

# Machine Learning Foundation HW1

B08502041 李芸芳

October 2020

## 第一題

答案：(D)

透過 supervised learning，先給機器幾張有標示品質好壞的芒果圖片，機器可以透過影像辨識猜測圖片大概長怎樣叫做好/不好的芒果，接下來就可以透過芒果的圖片判斷芒果的品質了。

## 第二題

答案：(E)

- (A) 單純由運氣決定是否為 spam。
- (B) 給人類決定是否為 spam。
- (C) 其實也是給人類決定是否為 spam，跟 (B) 的差別是 (C) 人類是給一個標準，判斷的工作交給機器，而 (B) 是人類自行設定標準接著判斷。
- (D) 機器沒有在學習，就只是在做計算與判斷而已。

## 第三題

答案：(D)

根據 Lecture 2 Page 16，更新次數  $T$  有上限  $\frac{R^2}{\rho^2}$ ，這個上限就是 worst case，因為  $R$  與  $\rho$  中  $\|\mathbf{x}_{n(t)}\|$  的次數一樣，所以上下抵銷，速度不變。

## 第四題

答案：(A)

根據推導， $R' = \max_n \|\eta_t \mathbf{x}_{n(t)}\|$ ， $\rho' = \min_n \frac{\eta_t y_{n(t)} \mathbf{w}_f^T \mathbf{x}_{n(t)}}{\|\mathbf{w}_f^T\|}$ ，更新次數  $T' \leq \frac{R'^2}{\rho'^2}$ ，得  $\frac{R'^2}{\rho'^2} = \frac{R^2}{\rho^2}$ ，所以  $p = 0$ 。

## 第五題

答案：(D)

$$\begin{aligned} y_{n(t)} \mathbf{w}_{t+1}^T \mathbf{x}_{n(t)} &= y_{n(t)} (\mathbf{w}_t + \eta_t y_{n(t)} \mathbf{x}_{n(t)})^T \mathbf{x}_{n(t)} \\ &= y_{n(t)} \mathbf{w}_t^T \mathbf{x}_{n(t)} + y_{n(t)}^2 \eta_t \|\mathbf{x}_{n(t)}\|^2 \\ &> 0 \end{aligned}$$

$$\begin{aligned} \because y_{n(t)}^2 &= 1 \therefore \eta_t \|\mathbf{x}_{n(t)}\|^2 > -y_{n(t)} \mathbf{w}_t^T \mathbf{x}_{n(t)} \\ \therefore \eta_t &> \frac{-y_{n(t)} \mathbf{w}_t^T \mathbf{x}_{n(t)}}{\|\mathbf{x}_{n(t)}\|^2} \end{aligned}$$

## 第六題

答案：(A)

令  $\lfloor \frac{-y_{n(t)} \mathbf{w}_t^T \mathbf{x}_{n(t)}}{\|\mathbf{x}_{n(t)}\|^2} + 1 \rfloor$  為  $\eta_t$ ，perfect line 的法向量為  $\mathbf{w}_f$ 。

$$\begin{aligned} \mathbf{w}_f^T \mathbf{w}_T &= (\mathbf{w}_t + \eta_t y_{n(t)} \mathbf{x}_{n(t)})^T \mathbf{w}_f \\ &= \mathbf{w}_f^T \mathbf{w}_{T-1} + \eta_t y_{n(t)} \mathbf{w}_f^T \mathbf{x}_{n(t)} \\ &\geq \mathbf{w}_f^T \mathbf{w}_{T-1} + \min_n \eta_t y_{n(t)} \mathbf{w}_f^T \mathbf{x}_{n(t)} \\ &\geq \mathbf{w}_f^T \mathbf{w}_{T-2} + 2 \min_n \eta_t y_{n(t)} \mathbf{w}_f^T \mathbf{x}_{n(t)} \\ &\dots \\ &\geq \mathbf{w}_f^T \mathbf{w}_0 + T \min_n \eta_t y_{n(t)} \mathbf{w}_f^T \mathbf{x}_{n(t)} \end{aligned}$$

$$\begin{aligned} \|\mathbf{w}_T\|^2 &= \|\mathbf{w}_{T-1} + \eta_t y_{n(t)} \mathbf{x}_{n(t)}\|^2 \\ &= \|\mathbf{w}_{T-1}\|^2 + 2\eta_t y_{n(t)} \mathbf{w}_{T-1}^T \mathbf{x}_{n(t)} + \eta_t^2 y_{n(t)}^2 \|\mathbf{x}_{n(t)}\|^2 \\ &\leq \|\mathbf{w}_{T-1}\|^2 + \eta_t^2 \|\mathbf{x}_{n(t)}\|^2 \\ &\leq \|\mathbf{w}_{T-1}\|^2 + \max_n \eta_t^2 \|\mathbf{x}_{n(t)}\|^2 \\ &\leq \|\mathbf{w}_{T-2}\|^2 + 2 \max_n \eta_t^2 \|\mathbf{x}_{n(t)}\|^2 \\ &\dots \\ &\leq \|\mathbf{w}_0\|^2 + T \max_n \eta_t^2 \|\mathbf{x}_{n(t)}\|^2 \end{aligned}$$

$$\begin{aligned} \text{設 } \rho &= \min_n \frac{\eta_t y_{n(t)} \mathbf{w}_f^T \mathbf{x}_{n(t)}}{\|\mathbf{w}_f\|}, R = \max_n \eta_t^2 \|\mathbf{x}_{n(t)}\|^2 \\ \therefore \mathbf{w}_0 &\text{ 可以為 } 0 \\ \therefore \frac{\mathbf{w}_f^T \mathbf{w}_T}{\|\mathbf{w}_T\| \|\mathbf{w}_f\|} &\geq \frac{T\rho}{\sqrt{TR}} = \frac{\sqrt{T}\rho}{R} \Rightarrow T \leq \frac{R^2}{\rho^2} \Rightarrow \text{保證在有完美的線時停止。} \end{aligned}$$

## 第七題

答案：(E)

機器透過回饋會大概知道對於一個盤面怎麼樣的玩法是好/不好的，進而進行修正，所以是 reinforcement learning。

## 第八題

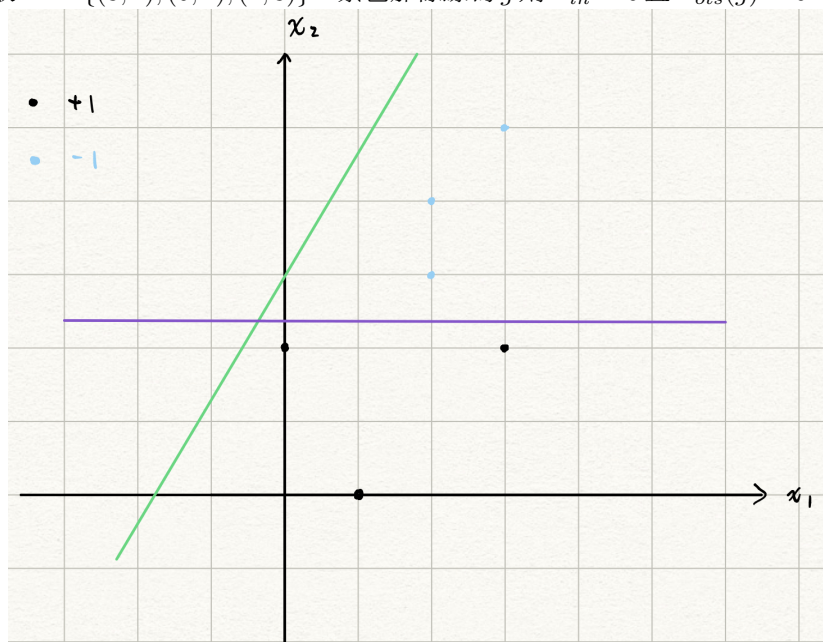
答案：(B)

給予部分影片告訴機器人類會怎麼做，接下來給予其他未標記的影片給機器自己學習是 batch learning 和 semi-supervised learning。影像需要透過機器轉換成對機器有用的資訊，所以影像是 raw feature。而每筆影像對應到的  $y_n$  是一個結構，所以是 structured learning。

## 第九題

答案：(E)

若取  $D = \{(1, 0), (3, 2), (0, 2)\}$ ，綠色那條線為  $g$  則  $E_{in} = 0$  但  $E_{ots}(g) = 1$ ；  
若取  $D = \{(3, 2), (0, 2), (2, 3)\}$ ，紫色那條線為  $g$  則  $E_{in} = 0$  且  $E_{ots}(g) = 0$ 。



## 第十題

答案：(B)

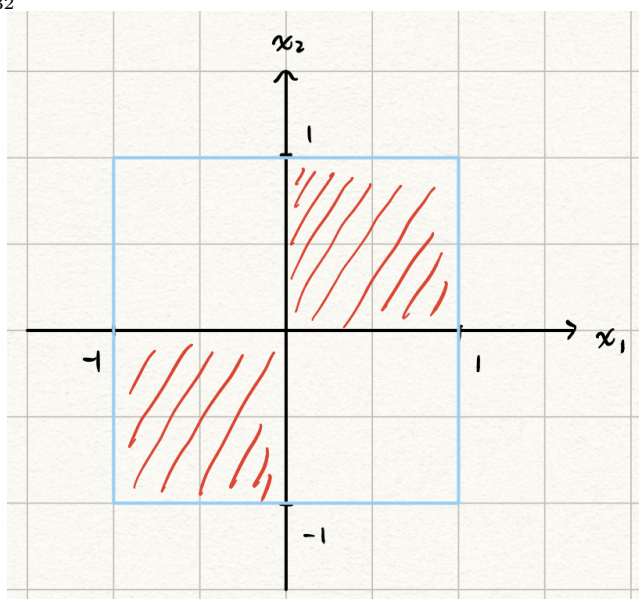
對的機率是  $1 - \delta \Rightarrow$  錯的機率是  $\delta$ 。而一個公正的硬幣每一面機率都應該是  $\frac{1}{2}$ ，但是題目偏差了  $\epsilon$ 。所以根據 Hoeffding Inequality

$$\delta = 2e^{-2\epsilon^2 N} \therefore N = \frac{1}{2\epsilon^2} \ln \frac{2}{\delta}$$

## 第十一題

答案：(C)

若取到圖中紅色部分則  $E_{in}(h_2) = 0$ ，紅色部分占全部  $\frac{1}{2}$ ，取 5 次則機率為  $(\frac{1}{2})^5 = \frac{1}{32}$ 。



## 第十二題

答案：(D)

根據下圖(紫色： $h_2(x) \neq f(x)$ 、黃色： $h_1(x) \neq f(x)$ 、綠色： $h_1(x) = h_2(x) = f(x)$ )可以發現並不存在  $h_1(x) \neq f(x)$  且  $h_2(x) \neq f(x)$  的狀況，因此若  $E_{in}(h_1)$  要等於  $E_{in}(h_2)$  且只取 5 個點，最多只會有三種情況： $E_{in}(h_1) = E_{in}(h_2) = 0, \frac{1}{5}, \frac{2}{5}$ 。

三種情況計算如下：

$$E_{in}(h_1) = E_{in}(h_2) = 0 \Rightarrow \left(\frac{3}{8}\right)^5 \times \frac{5!}{5!} = \frac{243}{32768}$$

(只取綠色區域)

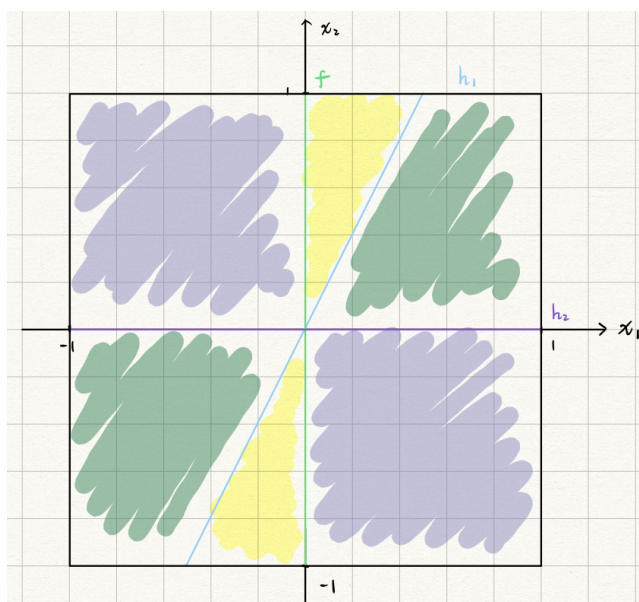
$$E_{in}(h_1) = E_{in}(h_2) = \frac{1}{5} \Rightarrow \left(\frac{3}{8}\right)^3 \times \left(\frac{1}{2}\right)^1 \times \left(\frac{1}{8}\right)^1 \times \frac{5!}{3!1!1!} = \frac{2160}{32768}$$

(綠色區域取三個點、紫色和黃色區域各取一個點，這樣  $E_{in}(h_1)$  和  $E_{in}(h_2)$  就會各錯一個點)

$$E_{in}(h_1) = E_{in}(h_2) = \frac{2}{5} \Rightarrow \left(\frac{3}{8}\right)^1 \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{8}\right)^2 \times \frac{5!}{2!2!1!} = \frac{1440}{32768}$$

(綠色區域取一個點、紫色和黃色區域各取兩個點，這樣  $E_{in}(h_1)$  和  $E_{in}(h_2)$  就會各錯兩個點)

全部加起來就會等於  $\frac{3843}{32768}$



### 第十三題

答案：(B)

對於  $i = 1, \dots, d$ ，若  $h_i$  的  $E_{in}$  是  $x$ ， $E_{out}$  是  $y$ ，則  $h_{i+d}$  的  $E_{in}$  是  $1 - x$ ， $E_{out}$  是  $1 - y$ 。

若有一筆資料對於  $h_i$  來說是 *Bad Data*，表示  $|x - y|$  很大，對於  $h_{i+d}$  來說也會是 *Bad Data*： $|E_{in} - E_{out}| = |(1 - x) - (1 - y)| = |y - x| = |x - y|$ 。

所以  $C$  可以取  $d$  就好。

## 第十四題

答案：(D)

可以得到5個綠3代表只能取  $BD$ ，因此機率為  $\frac{2^5}{4^5}$ 。

(A) 沒有這種情況，機率為  $\frac{0}{4^5}$

(B) 只能取  $C$ ，機率為  $\frac{1^5}{4^5}$

(C) 只能取  $ABD$ ，機率為  $\frac{3^5}{4^5}$

(D) 只能取  $AB$ ，機率為  $\frac{2^5}{4^5}$

(E) 只能取  $D$ ，機率為  $\frac{1^5}{4^5}$

故答案為(D)。

## 第十五題

答案：(C)

列舉對於每個數字要拿到5個綠色的情況。

5個綠1：無

5個綠2： $ABD$

5個綠3： $BD$

5個綠4： $AB$

5個綠5： $D$

5個綠6： $AC$

所以取法總共有  $(\text{全 } A + \text{全 } B + \text{全 } C + \text{全 } D) + (AC + AB + AD + BD)$   
+  $ABD$

$$A + B + C + D = 4$$

$$AC + AB + AD + BD = 4 \times (2^5 - 1 - 1)$$

$$ABD = \frac{3!}{2!} \frac{5!}{3!} + \frac{3!}{2!} \frac{5!}{2!2!}$$

( $\because ABD$  可以有  $(1, 1, 3)$  和  $(1, 2, 2)$  兩種取法)

$\therefore$  取法總共有 274 種。