

第一題

[←](#) 作業一
評分期限：40 min

截止時間 10月27日 23:59 PDT

 恭喜！您通過了！
通過條件 75% 或更高

堅持學習

成績
100%

作業一

最新提交作業的評分
100%

1. Which of the following problems are best suited for machine learning?
(i) Classifying numbers into primes and non-primes
(ii) Detecting potential fraud in credit card charges
(iii) Determining the time it would take a falling object to hit the ground
(iv) Determining the optimal cycle for traffic lights in a busy intersection
(v) Determining the age at which a particular medical test is recommended

10/10 分

 Correct

2. For Questions 2-5, identify the best type of learning that can be used to solve each task below.

10/10 分

 收到成績
通過條件 75% 或更高

成績
100%

查看反饋
我們會保留您的最高分數

第二題

利用半監督式學習，判斷一個人是不是快樂阿宅。首先先給幾個資料上標籤，人工判斷這幾筆資料的人是否為肥宅。資料變數包括一季追幾部動漫；手機、筆電桌布是不是動漫角色或場景；會不會喊我婆；懂不懂各種呼吸法，能把任何生物後面都加上呼吸；可以用美少女戰士出場順序來背日文的星期一到星期五。每個變數配上不同的加權，接著用這些標籤去把其他資料做個分類，我們就能獲得一個尋找快樂阿宅的半監督式學習法了。

第三題

$$\begin{aligned}
 E_f \{E_{OTS}(A(D), f)\} &= E_f \left\{ \frac{1}{L} \times \sum_{l=1}^L \mathbb{I}[A(D)(x_{N+l}) \neq f(x_{N+l})] \right\} \\
 &= E_f \left\{ \frac{1}{L} \times \sum_{l=1}^L l \times \binom{L}{l} \right\} \\
 &= \frac{1}{2^L} \times \left(\frac{1}{L} \times L \times 2^{L-1} \right) = \frac{1}{2}
 \end{aligned}$$

Therefore, $E_f = \text{constant}$ is proved

第四題

$$P = \left(\frac{1}{2}\right)^5 = \frac{1}{32} \quad (\text{the probability of picking bag A or D is } \frac{1}{2})$$

第五題

$$\begin{aligned}
 P &= P_{13} + P_2 + P_{46} + P_5 - P_{123} - P_{456} - P_{135} - P_{246} \\
 &= \underbrace{\left(\frac{1}{4} + \frac{1}{4}\right)^5}_{A \quad D} + \underbrace{\left(\frac{1}{4} + \frac{1}{4}\right)^5}_{B \quad D} + \underbrace{\left(\frac{1}{4} + \frac{1}{4}\right)^5}_{C \quad A} + \underbrace{\left(\frac{1}{4} + \frac{1}{4}\right)^5}_{B \quad C} - \underbrace{\left(\frac{1}{4}\right)^5}_D - \underbrace{\left(\frac{1}{4}\right)^5}_C - \underbrace{\left(\frac{1}{4}\right)^5}_A - \underbrace{\left(\frac{1}{4}\right)^5}_B \\
 &= \frac{31}{256}
 \end{aligned}$$

(P_{index} : $index$ 為綠色面點數的機率，例如 P_{13} 代表五個骰子點數 1 和 3 同時是綠色的機率)

比較答案後會發現第四題其實就是第五題的其中一個 case 而已(P_{13})。直觀來

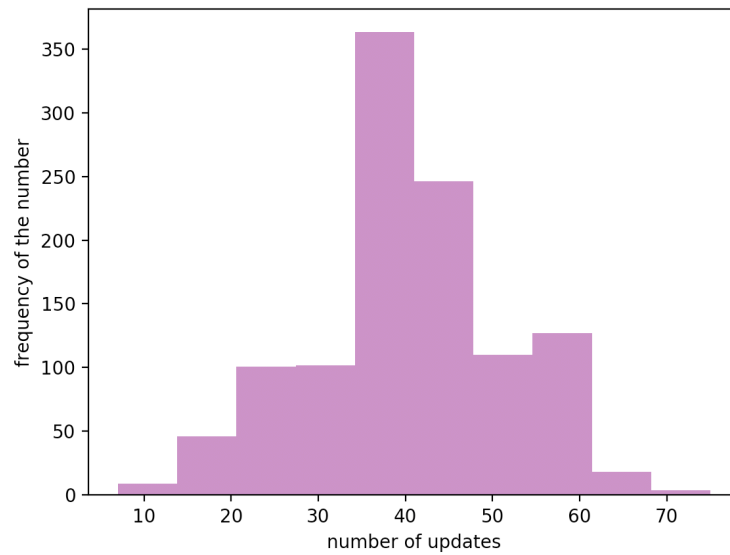
看，第五題的答案應該為第四題的六倍，因為數字從 1 變成了 some number，

但觀察 conditions 後又會直觀地認為答案應該是四倍，因為 1 和 3 會同時出

現，4 和 6 會同時出現。但因為有 over-lapping 的情況，所以實際答案更少。

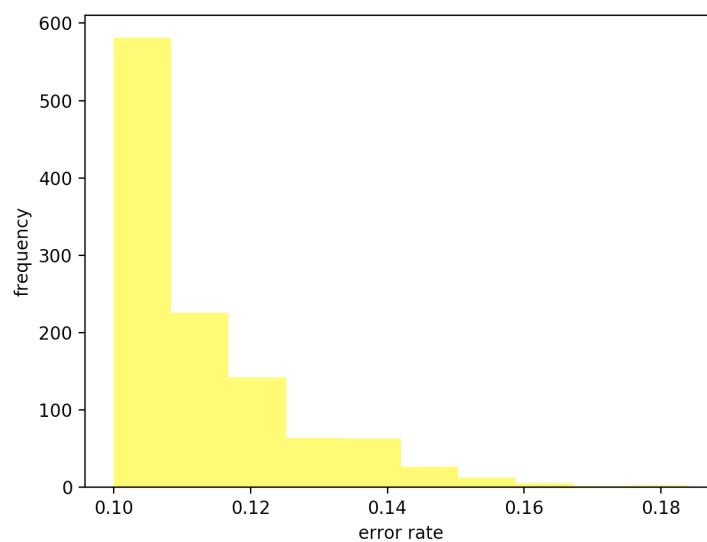
第六題

Average number of update = 40.151865008880996



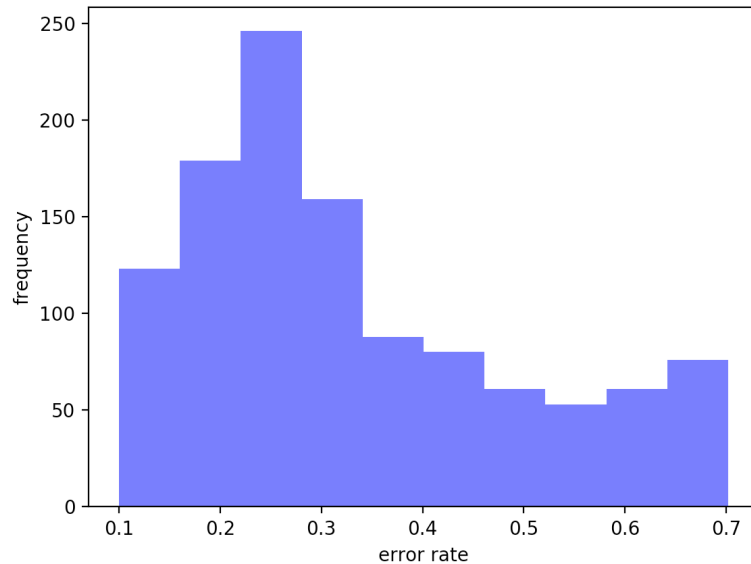
第七題

Average error rate = 0.11287566607459974



第八題

Average error rate = 0.33847779751332135



可以看出 error rate 上升不少，從 0.113 左右上升至 0.338 左右，表示使用 Pocket 來做 training 會比一般的 PLA 要好。其一是 Pocket 可避免資料無法線性可分的情況，其二是錯誤率會修正的更低。

第九題

這方法不可行。因為從投影片十五我們可以知道：

start from $\mathbf{w}_0 = \mathbf{0}$, after T mistake corrections,

$$\frac{\mathbf{w}_f^T}{\|\mathbf{w}_f\|} \frac{\mathbf{w}_T}{\|\mathbf{w}_T\|} \geq \sqrt{T} \cdot \text{constant}$$

時間與 T 成正比。又從投影片十六可得 $T \leq \frac{R^2}{\rho^2}$ ，其中

$$R^2 = \max_n \|\mathbf{x}_n\|^2 \quad \rho = \min_n y_n \frac{\mathbf{w}_f^T}{\|\mathbf{w}_f\|} \mathbf{x}_n$$

據題意，現在把每個 \mathbf{x} 都除以十倍，不難發現與時間有關的關係式已經把 \mathbf{x}_n 這項變數消掉了，因此就算對 \mathbf{x}_n 這項變數去做操作，也不會影響 PLA 演算法跑的時間