

1.

項目	狀態	截止時間	權重	成績
 作業一 測驗	已通過	11月18日 14:59 CST	50%	100%

2.

網頁分類，將給定的網頁分類至數個類別中的其中一個，例如:教育、娛樂、線上購物...等類別。由於對大量的網頁進行人工標註是相當耗時費力的，故應採用 semi-supervised learning 的方式，只對部分的樣本進行標註，來降低資料取得的成本。

3.

$E_f\{E_{OTS}(A(D), f)\} = \text{constant}$  為 True

Proof:

由於在 test set 中有  $L$  個 example, 故可能的答案有  $L$  種

對於任意  $A(D)$ , 這  $L$  種中有

$C_0$  種與  $A(D)$  的答案相同 (全相同)

$C_1$  種與  $A(D)$  的答案相差 1 個

$\vdots$

$C_k$  種與  $A(D)$  的答案相差  $k$  個

$\vdots$

$C_L$  種與  $A(D)$  的答案相差  $L$  個 (全相異)

故  $E_f\{E_{OTS}(A(D), f)\} = \frac{1}{L}(C_0 * 0 + C_1 * 1 + C_2 * 2 + \dots + C_L * L)$

為 - constant

## 4.

由於只有 A,D 具有 green 1's,又每次抽到 A 或 D 的機率是  $2/4$

故答案為  $(1/2)^5$

也就是  $1/32$

## 5.

若要使所有的 x 為 green,則 5 次都只能抽到:y

x	y
1	(A,D)
2	(B,D)
3	(A,D)
4	(B,C)
5	(A,C)
6	(B,C)

5 次內都只抽到(A,D)的機率為  $(1/2)^5=1/32$  ((B,D),(B,C),(A,C)同理, 皆為  $1/32$ )

然而都只抽到(A,D)和(A,C)的狀況都包含了只抽到(A)的狀況

只抽到(B,D)和(B,C)的狀況都包含了只抽到(B)的狀況

只抽到(B,C)和(A,C)的狀況都包含了只抽到(C)的狀況

只抽到(A,D)和(B,D)的狀況都包含了只抽到(D)的狀況

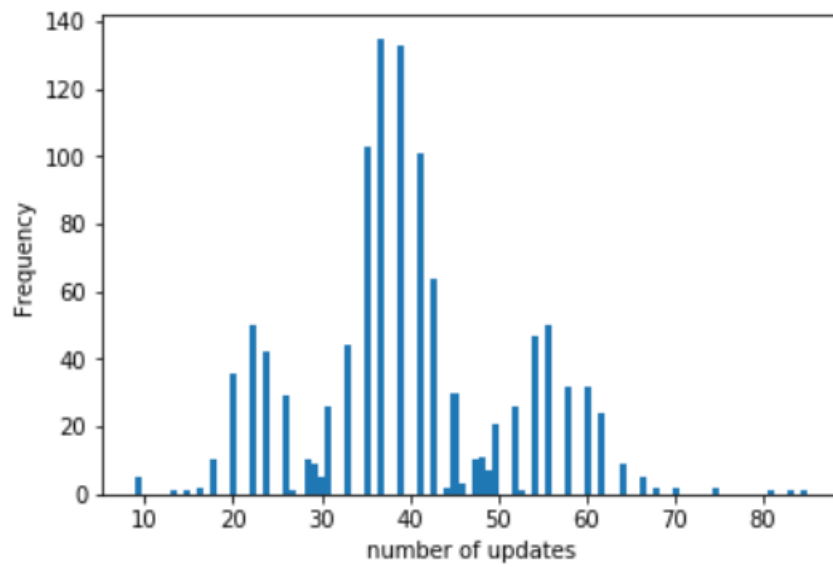
又只抽到(A)的機率為  $(1/4)^5=1/1024$  ((B),(C),(D)同理, 皆為  $1/1024$ )

故結果為  $(1/32)^4 - (1/1024)^4 = 31/256$

比起限定 1 為全綠的狀況(第 4 題), 本題不限定全綠的數字, 故機率較大, 然而不是變成原本的 6 倍, 因為數字 1~6 全綠的條件有部分重疊

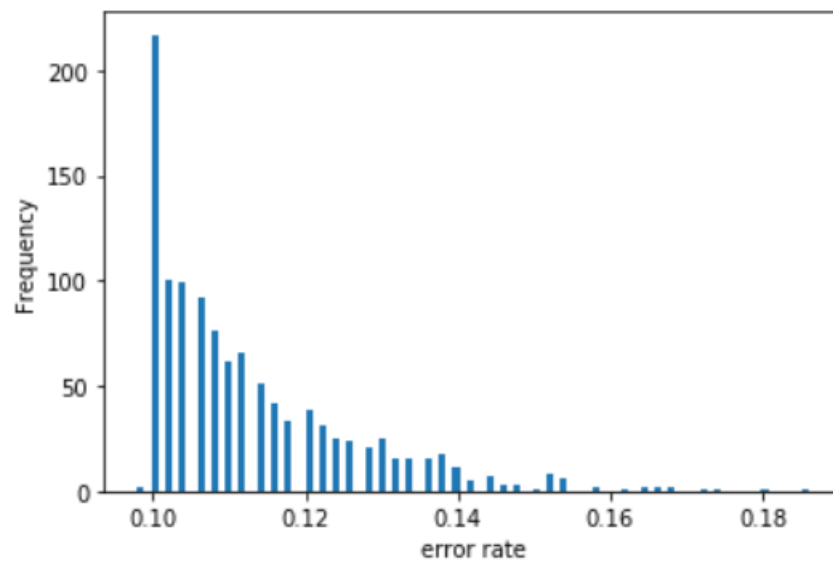
6.

average number of updates before the algorithm halts: about 40.0337



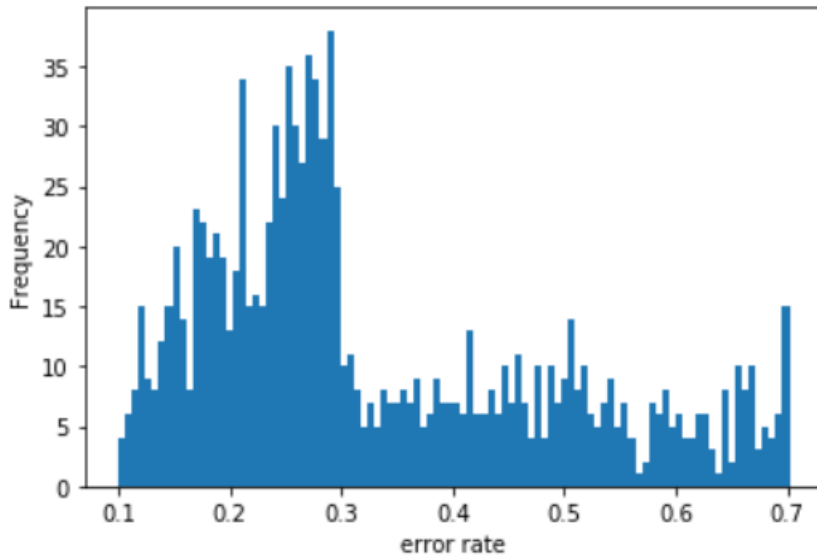
7.

average error rate: about 0.1127



8.

average error rate: about 0.3316



本題是更新到第 100 次所得的結果，而[7]則是更新 100 次中最好的一次，故[7]的 error rate 理所當然會比本題好

9.

根據第三週的課程投影片(9/7)的第30頁

$T \leq \frac{R^2}{\rho^2}$  , 其中  $T$  為 mistake corrections 次數

$$R^2 = \max_n \|x_n\|^2$$

$$\rho = \min_n y_n \frac{w_s^T}{\|w_s\|} x_n$$

$$\Rightarrow T \leq \frac{\max_n \|x_n\|^2}{\left(\min_n y_n \frac{w_s^T}{\|w_s\|} x_n\right)^2}$$

$$\text{令 } x_n' = \frac{1}{10} x_n \Rightarrow \|x_n'\| = \frac{1}{10} \|x_n\|$$

$$\Rightarrow T \leq \frac{\max_n \|x_n'\|^2}{\left(\min_n y_n \frac{w_s^T}{\|w_s\|} x_n'\right)^2} = \frac{\max_n \cancel{\frac{1}{10^2}} \|x_n\|^2}{\cancel{\frac{1}{10^2}} \left(\min_n y_n \frac{w_s^T}{\|w_s\|} x_n\right)^2} = \frac{\max_n \|x_n\|^2}{\left(\min_n y_n \frac{w_s^T}{\|w_s\|} x_n\right)^2}$$

結論: 對  $x_n$  進行線性縮小無法獲得更快的效果