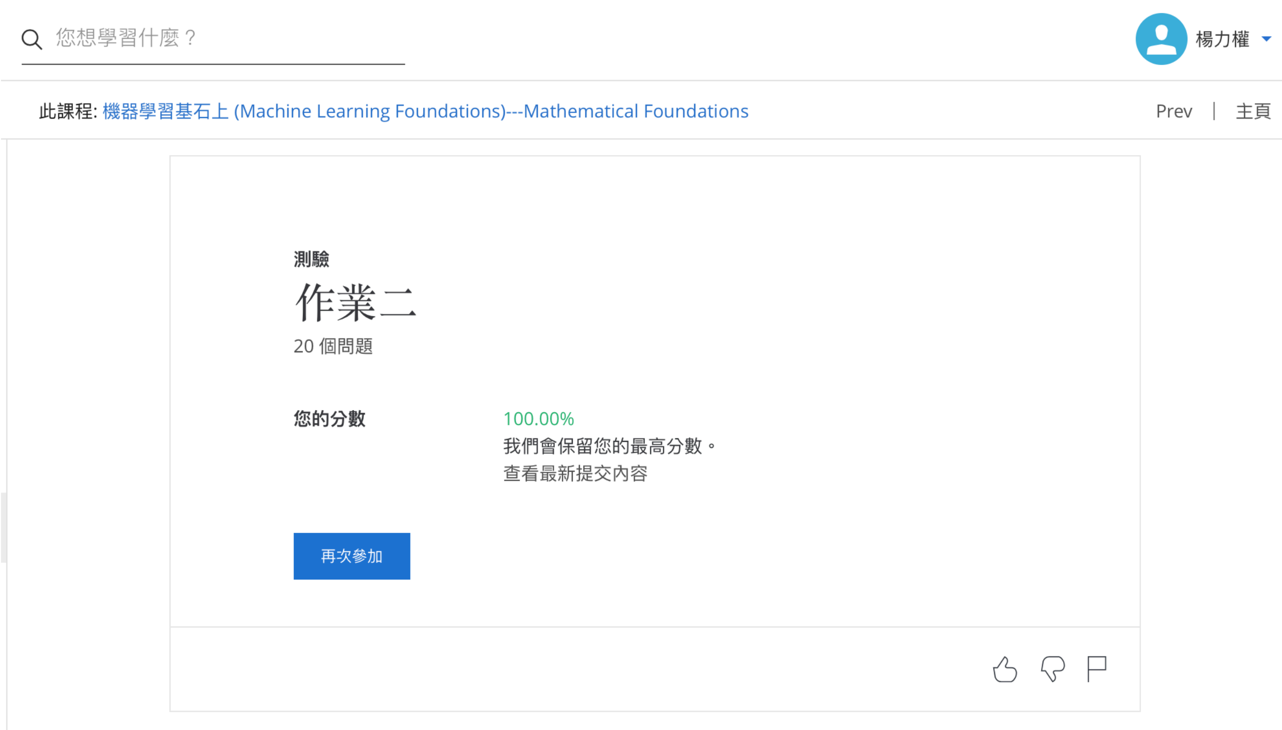
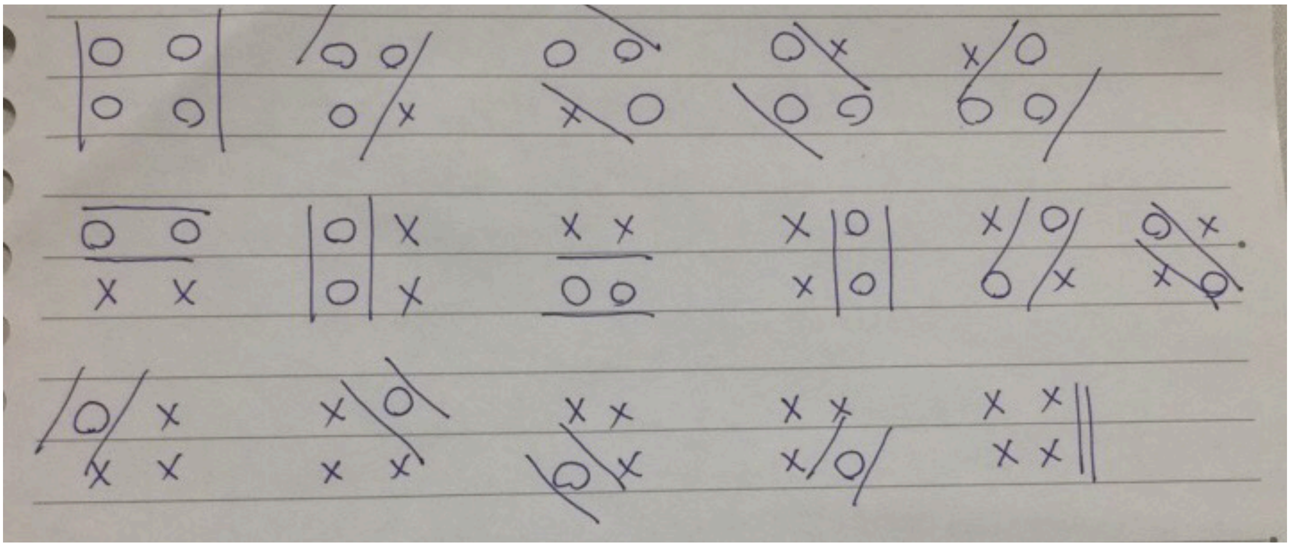
MLFoundation hw2

r07922100 楊力權

1.



2.



N=4的時候thick line可以shatter 2^4=16種dichotomy，因此$v_{dc} \geq 4$也就是說thick line的vc dimension不小於4。

3.

取點$x \in \{4^i | i \in Z^+\}$，$\alpha = \sum_{i=1}^{N} (\frac{1}{4})^i a_i$，$a_i \in \{0,2\}$，討論N的情況如下：

N=1：$x_1 = 4,\ \alpha = \frac{1}{4}a_1$，則$h_\alpha(x_1) = sign(|a_1 \% 4 - 2| - 1)$，因此調整$a_1 = 0$可以得到$h_\alpha(x_1) = +1$，調整$a_1 = 2$可以得到$h_\alpha(x_1) = -1$，**N=1可以shatter**。

N=2：$x_1 = 4,\ x_2 = 16,\ \alpha = \frac{1}{4}a_1 + \frac{1}{16}a_2$，$h_\alpha(x) = sign(|(a_1 + \frac{1}{4}a_2)x \% 4 - 2| - 1)$

	$y_2 = +1, a_2 = 0$	$y_2 = -1, a_2 = 2$
$y_1 = +1$ $a_1 = 0$	$h_\alpha(x_1) = sign(\|(0+0) \% 4 - 2\| - 1) = +1$ $h_\alpha(x_2) = sign(\|(0+0) \% 4 - 2\| - 1) = +1$	$h_\alpha(x_1) = sign(\|(\frac{2}{16}4) \% 4 - 2\| - 1) = +1$ $h_\alpha(x_2) = sign(\|(\frac{2}{16}16) \% 4 - 2\| - 1) = -1$
$y_1 = -1$ $a_1 = 2$	$h_\alpha(x_1) = sign(\|(\frac{2}{4}4) \% 4 - 2\| - 1) = -1$ $h_\alpha(x_2) = sign(\|(\frac{2}{4}16) \% 4 - 2\| - 1) = +1$	$h_\alpha(x_1) = sign(\|(\frac{2}{4} + \frac{2}{16})4 \% 4 - 2\| - 1) = -1$ $h_\alpha(x_1) = sign(\|(\frac{2}{4} + \frac{2}{16})16 \% 4 - 2\| - 1) = -1$

N=2可以shatter。

假設N=k的時候可以shatter，則加入第k+1個點(N=k+1)$a_{k+1} = 4^{k+1}$，討論加入的點k+1以及加入點k+1後任一點m的情形(0<m<k+1)：

$$h_\alpha(x_N) = sign(|(\sum_{i=1}^{k+1} \frac{1}{4^i} a_i)4^{k+1} \% 4 - 2| - 1) = sign(|(a_{k+1} + 4a_k + \ldots + 4^k a_1) \% 4 - 2| - 1)$$

$$= sign(|(a_{k+1}) \% 4 - 2| - 1)$$

$$h_\alpha(x_m) = sign(|(\sum_{i=1}^{k+1} \frac{1}{4^i} a_i)4^m \% 4 - 2| - 1)$$

$$= sign(|(\frac{1}{4^{m-k-1}} a_{k+1} + \ldots + \frac{1}{4} a_{m+1} + a_m + 4a_{m-1} + \ldots + 4^{m-1} a_1) \% 4 - 2| - 1)$$

因$a_m \in \{0,2\}$所以若$0 \le (\frac{1}{4^{m-k-1}} a_{k+1} + \ldots + \frac{1}{4} a_{m+1}) < 1$則

$$h_\alpha(x_m) = sign(|(\frac{1}{4^{m-k-1}} a_{k+1} + \ldots + \frac{1}{4} a_{m+1} + a_m) \% 4 - 2| - 1) = sign(|(a_m) \% 4 - 2| - 1)$$

而因$(\frac{1}{4^{m-k-1}} a_{k+1} + \ldots + \frac{1}{4} a_{m+1}) \le 2(\frac{1}{4} + \frac{1}{4^2} + \ldots) = 2\frac{\frac{1}{4}}{\frac{3}{4}} = \frac{2}{3} < 1$使上式成立。

因此任一點m(包括第k+1點)取$a_m = 0$則$h_\alpha(x_m) = +1$，取$a_m = 2$則$h_\alpha(x_m) = -1$，表示可以透過調整α內的參數$a_1 \ldots a_{k+1}$組合成一個可以shatter k+1個點的α，也就是說就算$N = \infty$也能夠shatter，因此triangle wave的$d_{vc} = \infty$。

4.

hypothesis H_1在data數量$N = d_{vc}(H_1) + 1$時不能被shatter

hypothesis H_2在data數量$M = d_{vc}(H_2) + 1$時不能被shatter

hypothesis $H_1 \cap H_2$必須同時符合H_1與H_2，在data量≥N與M必不能被shatter

所以break point $k \le min(N, M)$，$d_{vc}(H_1 \cap H_2) \le min(d_{vc}(H_1), d_{vc}(H_2)) \le d_{vc}(H_2)$

且若H_1完全包含H_2，則不等式等號成立，$d_{vc}(H_1 \cap H_2) = d_{vc}(H_2)$

5.

positive ray與negative ray的聯集是1D perceptron，$m_{H_1 \cup H_2}(N) = 2N$

$m_{H_1 \cup H_2}(1) = 2 = 2^1$ $m_{H_1 \cup H_2}(2) = 4 = 2^2$ $m_{H_1 \cup H_2}(3) = 6 \ne 2^3$

發現break point=3，因此$d_{vc}(H_1 \cup H_2) = 2$

6.

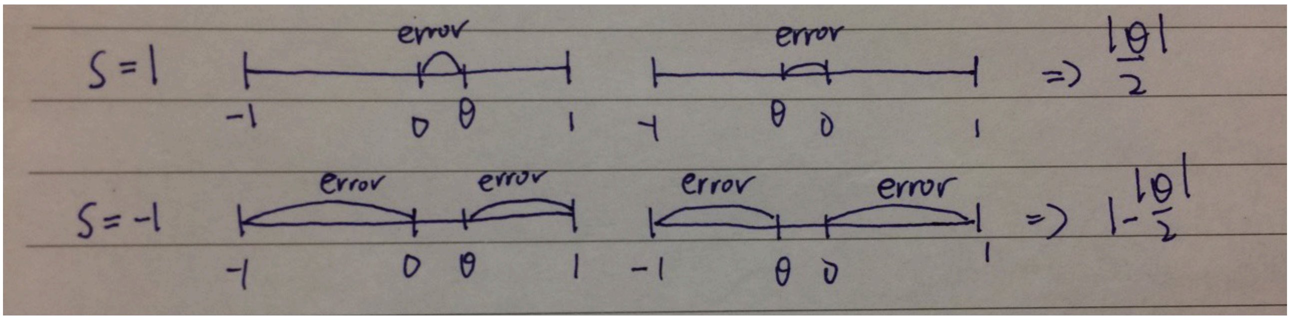
μ : hypothesis H錯誤率, λ : data沒有被noise flip的機率

整體的錯誤率，是H的錯誤率μ乘上沒有被noise干擾的機率λ，加上H正確率

$1-\mu$乘上被noise干擾而flip的機率$1-\lambda$，因此可得下列error rate式子。

錯誤率 $E = \mu\lambda + (1-\mu)(1-\lambda)$

討論s的兩種情形，+1與-1的錯誤率如下：



可得error rate $\mu = \begin{cases} \frac{|\theta|}{2} & \text{if } s = +1 \\ 1 - \frac{|\theta|}{2} & \text{if } s = -1 \end{cases} = (\frac{s+1}{2})(\frac{|\theta|}{2}) + (\frac{1-s}{2})(\frac{2-|\theta|}{2})$

$\Rightarrow \mu = \frac{s(|\theta|-1)+1}{2}$且$1-\mu = \frac{1-s(|\theta|-1)}{2}$

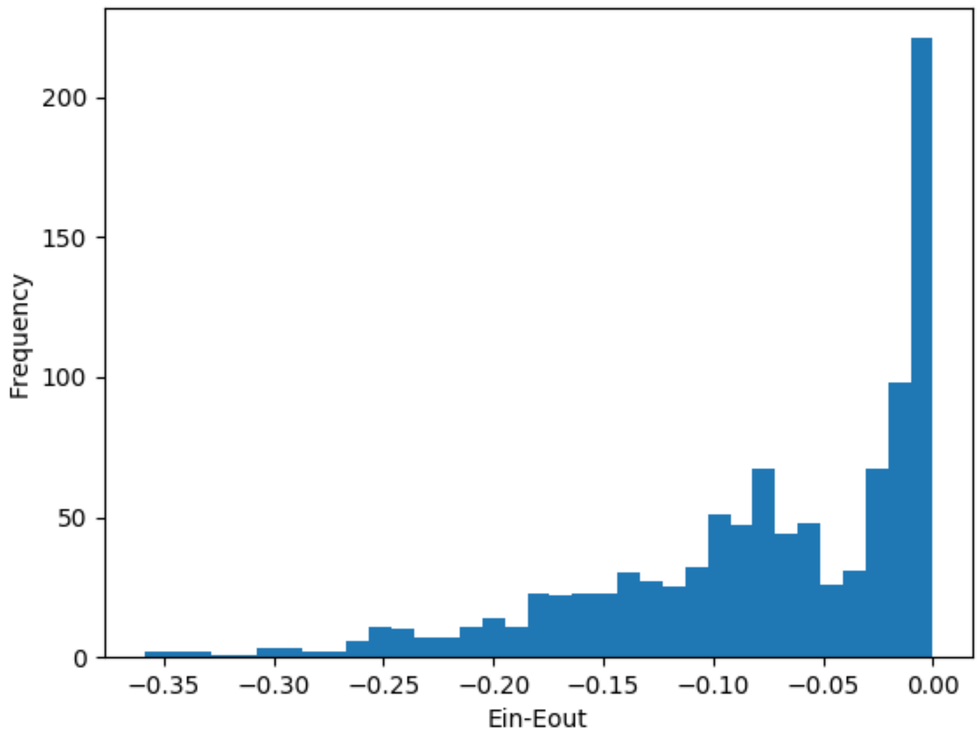
且noise rate 0.2可得 $\lambda = 0.8$，則可得

$E_{out}(h_s, \theta) = 0.8\frac{s(|\theta|-1)+1}{2} + 0.2\frac{1-s(|\theta|-1)}{2} = 0.5 + 0.3s(|\theta|-1)$

7.

平均E_in=0.1736，平均E_out=0.2514

機率分佈大致符合E_in接近E_out，也就是大部分的情況下，E_in與E_out是相差不遠的(較靠近圖右的0處)，儘管如此，差到0.3甚至0.35的特例還是存在，只是相對機率極低。



8.

上課老師有證明過$B(N,k) \leq \sum_{i=0}^{k-1} \binom{N}{i}$，而$\sum_{i=0}^{k-1} \binom{N}{i}$的意義其實可以視為如下：

在$x \in \{-1,1\}$內N個樣本中挑出$\binom{N}{0}$個為-1的dichotomy，加上$\binom{N}{1}$個為-1的dichotomy......直到加上$\binom{N}{k-1}$的dichotomy時共$\sum_{i=0}^{k-1} \binom{N}{i}$個dichotomy。

可以發現在這些dichotomy中任取k個點，必無法shatter因為至少要出現k個-1的情形，因此可知在break point=k的情況下$m_H(N)$至少$\sum_{i=0}^{k-1} \binom{N}{i}$個，可得

$$B(N,k) \geq \sum_{i=0}^{k-1} \binom{N}{i}$$

則$B(N,k) \leq \sum_{i=0}^{k-1} \binom{N}{i} \cap B(N,k) \geq \sum_{i=0}^{k-1} \binom{N}{i} \equiv B(N,k) = \sum_{i=0}^{k-1} \binom{N}{i}$