Machine Learning – Assignment 2

103062528 林玉山

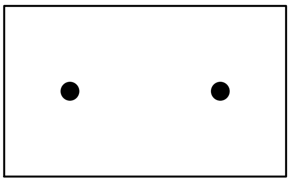
1. 證明長方形的在二維空間上的 VC Dimension 是 4？

基本上，只要能夠在 1~4 個點的情況下，各找到一種 placement，並且可以讓長方形成功讓所有不同 labels 的組合都能夠分開的話，就能證明其 VC Dimension 至少為 4。

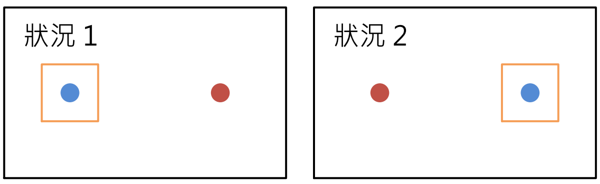
(以下標示，正的點為藍色，負的點為紅色，橘色線為 hypothesis)

A. 一個點，不需要分

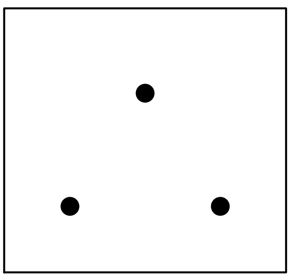
B. 兩個點，針對下面這種擺法：



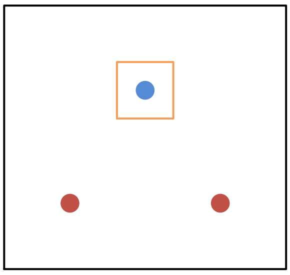
(a) 只有兩種情形，分別可以這樣分：



C. 三個點，針對下面這種擺法：

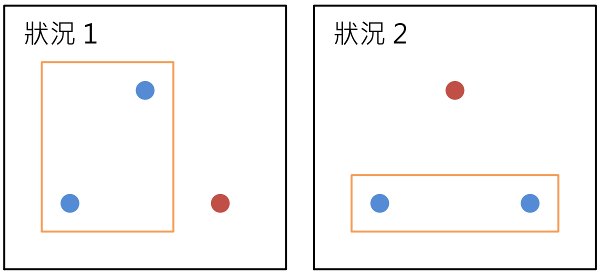


(a) 首先我們先看，只有一個點是正的，其他兩個都是負的情況。我們可以這樣切：



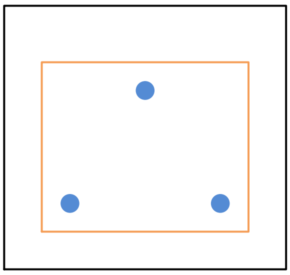
正的點換成左下或右下的點，也是一樣的切法。

(b) 如果有兩個正的點，一種是上面配下面任一點(只畫出與左邊的點配對的圖)，另一種是下面兩個點：



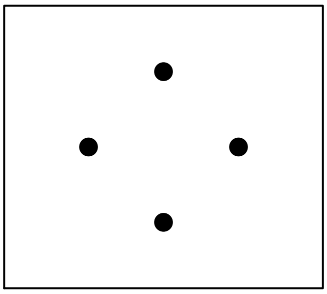
可見都可以切開

(c) 如果三個點都是正的：

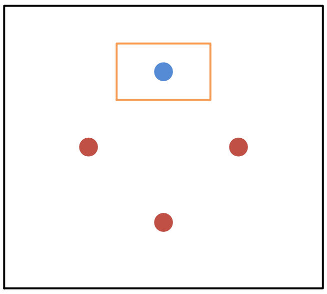


也明顯沒有問題，因此可以證明三個點在這種 placement 下，任意 labels 的配法都可以切開

D. 三個點，針對下面這種擺法：



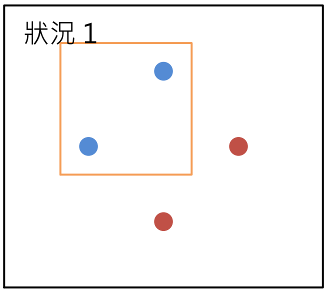
(a) 對於只有一個正，其他三點為負的情況：

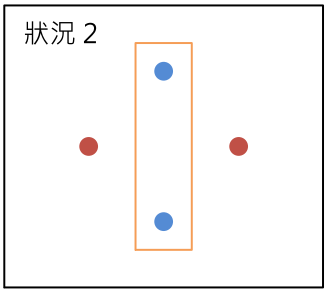


只有一個正的情形下，其他三個點分別為正的情形類似。

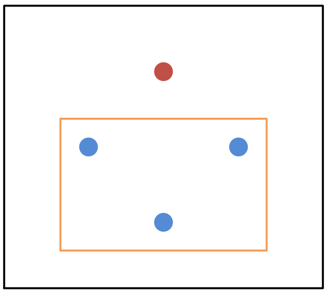
(b) 對於兩個正，兩個負的情形，有以下幾種狀況：

一種是相鄰兩點為正，一種是相對兩點為正，分別可以這樣切：





(c) 對於三個是正的，剩下一個是負的，只有這種情況：

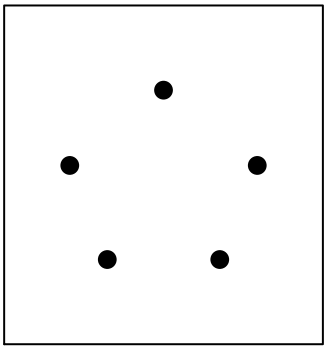


其他種狀況都類似。

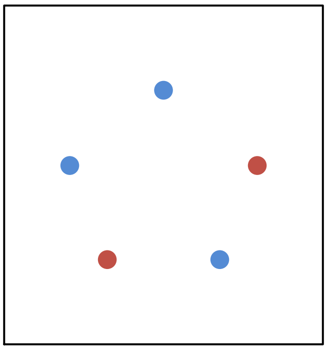
(d) 四個點都是正只要全部框進去即可。

E. 除了證明一到四個點都可行之外，還要證明五個點不可行。

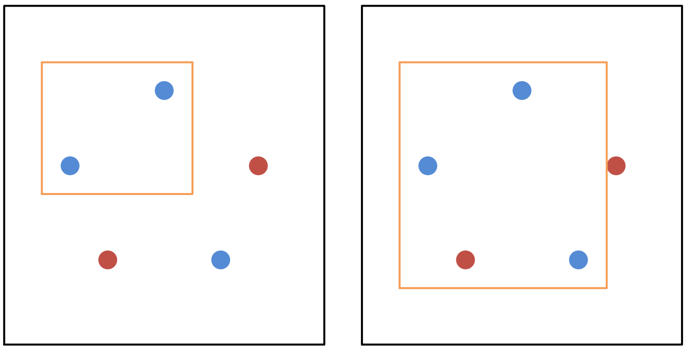
這邊舉出五個點最有可能被分開的情況，也就是五邊形的擺法：



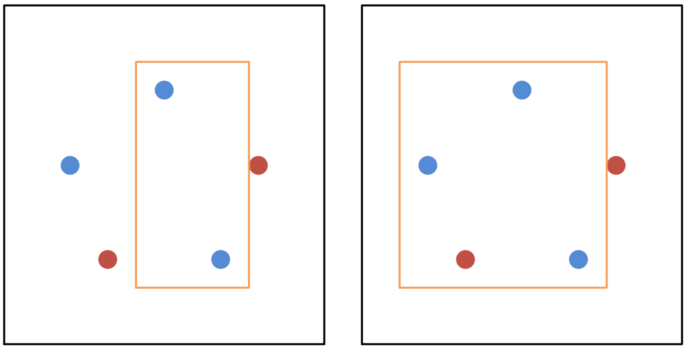
考慮這組 label：



如果我們先涵蓋左上角的兩個點，並拉伸長方形以納入右下角的點，就會變成這樣：



如果我先涵蓋右邊的兩個點，在往左拉伸，就會變成這樣：



只要我想要涵蓋左邊跟右下角的點，就一定會涵蓋到左下角的點。因為左邊的點比左下角的點更左邊，右下角的點跟左下角的點又在同一個水平。而這個 hypothesis 的長方形只有水平跟垂直的邊，不可能以斜線的方式呈現，所以五個點的狀況下不可能所有 case 都切的開。

因此我們証明了 1~4 個點，hypothesis 都能切開，但是 5 不行。因此這個 hypothesis 的 VC Dimension 為 4。

2. 證明下式成立

$$\boldsymbol{E\left[e^{\eta f(Z)}\right] \leq \frac{b - E[f(Z)]}{b - a} e^{\eta a} + \frac{E\big(f(Z)\big) - a}{b - a} e^{\eta b}}$$

因為 e^x 是 Convex Function，由 Convex 的性質我們可以得知

$$e^{(1-\theta)x+\theta y} \leq (1-\theta)e^x + \theta e^y, \theta \in [0,1]$$

令一變數為 $u \in [a,b]$ ，我們將 θ 代換為 $\frac{u-a}{b-a}$ ，因此上式變為

$$e^{\left(1-\frac{u-a}{b-a}\right)x+\frac{u-a}{b-a}y} \leq \left(1-\frac{u-a}{b-a}\right)e^x + \frac{u-a}{b-a}e^y$$

$$e^{\left(\frac{b-a-u+a}{b-a}\right)x+\frac{u-a}{b-a}y} \leq \frac{b-a-u+a}{b-a}e^x + \frac{u-a}{b-a}e^y$$

$$e^{\frac{b-u}{b-a}x+\frac{u-a}{b-a}y} \leq \frac{b-u}{b-a}e^x + \frac{u-a}{b-a}e^y$$

此時，我們再令 $x = \eta a, y = \eta b$

$$e^{\frac{b-u}{b-a}\eta a+\frac{u-a}{b-a}\eta b} \leq \frac{b-u}{b-a}e^{\eta a} + \frac{u-a}{b-a}e^{\eta b}$$

$$e^{\eta\frac{ba-ua+ub-ab}{b-a}} \leq \frac{b-u}{b-a}e^{\eta a} + \frac{u-a}{b-a}e^{\eta b}$$

$$e^{\eta\frac{ub-ua}{b-a}} \leq \frac{b-u}{b-a}e^{\eta a} + \frac{u-a}{b-a}e^{\eta b}$$

$$e^{\eta u} \leq \frac{b-u}{b-a}e^{\eta a} + \frac{u-a}{b-a}e^{\eta b}$$

再來，我們左右各取期望值

$$E[e^{\eta u}] \leq E\left[\frac{b-u}{b-a}e^{\eta a} + \frac{u-a}{b-a}e^{\eta b}\right]$$

根據絕對值的性質，並且 a, b, η 皆為常數

$$E[e^{\eta u}] \leq \frac{b-E[u]}{b-a}e^{\eta a} + \frac{E[u]-a}{b-a}e^{\eta b}$$

最後，我們令 $u = f(Z)$

$$E\left[e^{\eta f(Z)}\right] \leq \frac{b-E[uf(Z)]}{b-a}e^{\eta a} + \frac{E[f(Z)]-a}{b-a}e^{\eta b}$$

得證。

3. 證明如果 Sauer's Lemma 成立，則下式成立：

$$S_H(N) \le \left(\frac{eN}{v}\right)^v$$

證明如下：由 Sauer's Lemma，我們可以得到下式：

$$S_H(N) \le \sum_{k=0}^{v} \binom{N}{k}$$

因為 $N > v$，且 $k \le v$ 因此我們可以將上式寫成下式：

$$S_H(N) \le \sum_{k=0}^{v} \binom{N}{k} \le \sum_{k=0}^{v} \binom{N}{k} (\frac{N}{v})^{v-k}$$
$$S_H(N) \le \left(\frac{N}{v}\right)^v \sum_{k=0}^{v} \binom{N}{k} \left(\frac{v}{N}\right)^k$$

此時我們將上式做為 1 式。另外，根據 Binomial Formula

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

我們得知

$$\left(1+\frac{v}{N}\right)^N = \sum_{k=0}^{N} \binom{N}{k} 1^{N-k} (\frac{v}{N})^k = \sum_{k=0}^{N} \binom{N}{k} (\frac{v}{N})^k$$

因此 1 式可以寫成

$$S_H(N) \le \left(\frac{N}{v}\right)^v \left(1+\frac{v}{N}\right)^N$$

將上式作為 2 式。而根據 Hint，我們知道

$$e^v \ge \left(1+\frac{v}{N}\right)^N$$

因此 2 式又可以寫成

$$S_H(N) \le \left(\frac{N}{v}\right)^v \left(1+\frac{v}{N}\right)^N \le \left(\frac{N}{v}\right)^v e^v = \left(\frac{eN}{v}\right)^v$$

得證。

4. 證明

$$\boldsymbol{var}[\mathbf{1}(\boldsymbol{g}^*(\boldsymbol{x}) \neq \boldsymbol{r})] \leq \frac{\mathbf{1}}{\mathbf{4}}$$

我們先計算其 variance

$$\boldsymbol{v}ar[1(g^*(x) \neq r)] = E\left[\left(1(g^*(x) \neq r)\right)^2\right] - E[1(g^*(x) \neq r)]^2$$

$$= \left(1^2 \times p + 0^2 \times (1-p)\right) - \left(1 \times p + 0 \times (1-p)\right)^2$$

$$= p - p^2 = p(1-p)$$

此時我們使用算幾不等式，得到

$$\boldsymbol{v}ar[1(g^*(x) \neq r)] = \sqrt{p(1-p)} \leq \frac{p + (1-p)}{2}$$

$$\Rightarrow \sqrt{\boldsymbol{v}ar[1(g^*(x) \neq r)]} \leq \frac{1}{2}$$

$$\Rightarrow \boldsymbol{v}ar[1(g^*(x) \neq r)] \leq \frac{1}{4}$$

得證。