

Hw3

1.

QUIZ

作業三

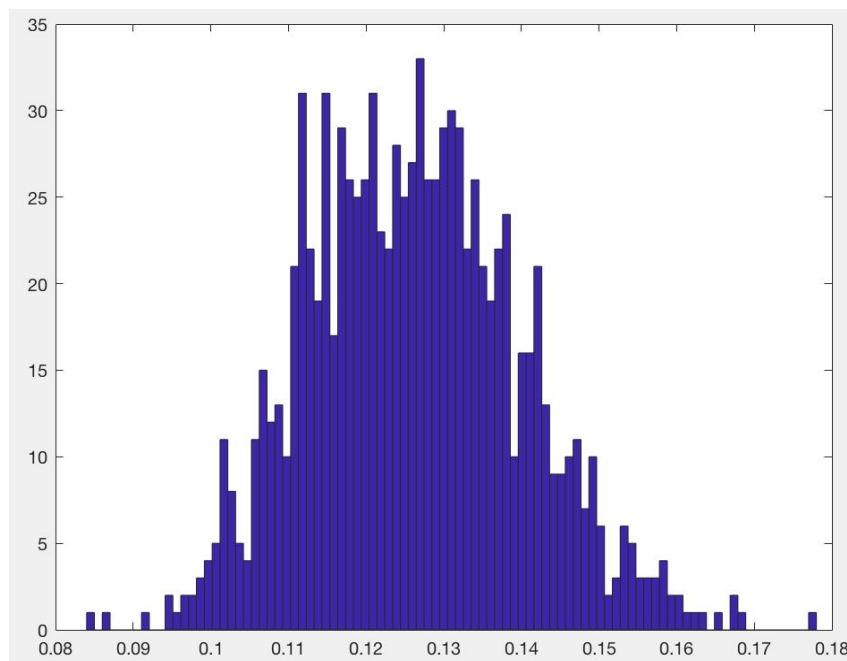
20 questions

Your Score 200/200 points (100%)
We keep your highest score.
[View Latest Submission](#)

[Take it again](#)

7.

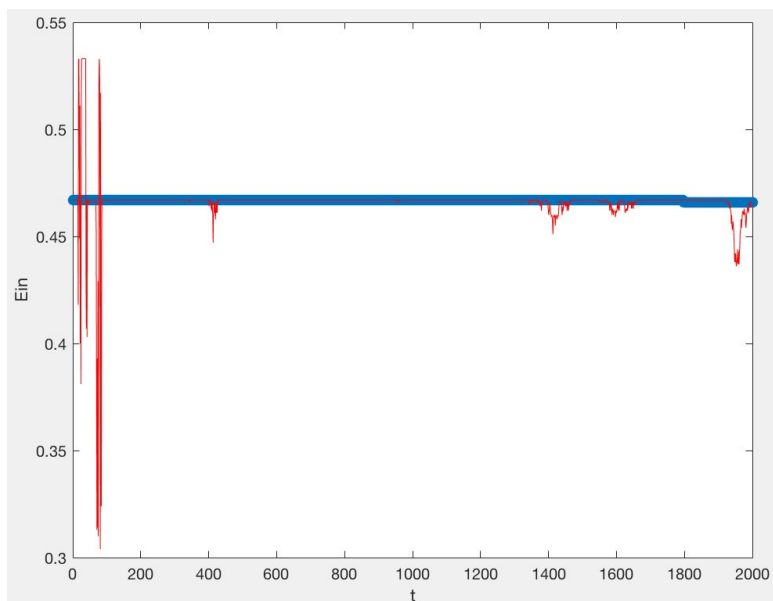
aver Eout : 0.126259



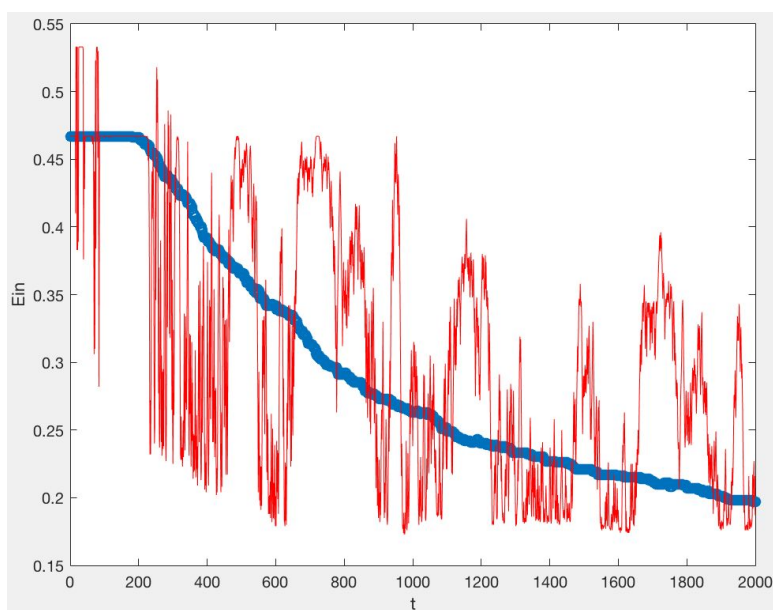
X軸代表Eout的值, Y軸代表各數值累積的次數
總共做了1000次, 平均Eout = 0.126

8.

```
Ein for gradientDescent with eta = 0.001 : 0.466000  
Ein for stochasticGradient with eta = 0.001 : 0.464000  
Ein for gradientDescent with eta = 0.01 : 0.197000  
Ein for stochasticGradient with eta = 0.01 : 0.187000
```



eta = 0.001



eta = 0.01

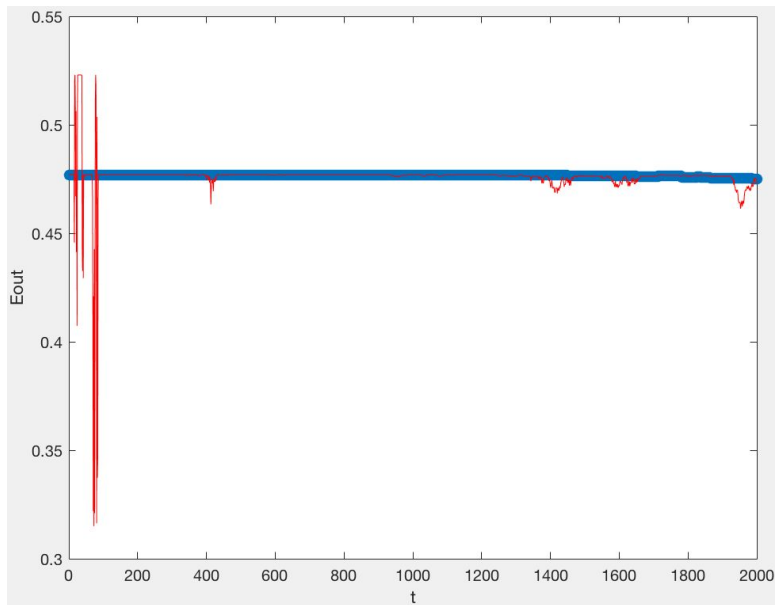
以上兩張圖的藍線代表the gradient descent version, 紅線代表the stochastic gradient descent version.

可以觀察出來, 如果eta設太小更新起來會太慢, 2000次的更新做不出什麼效果, 而eta=0.01的話不管是哪一種方法最後 E_{in} 都會小於0.2, 更新起來就有效果. 再來是原本的gradient descent因為會看所有的點, 所以做起來比較穩定, 在eta=0.01的圖上有穩定的下降; 而stochastic gradient descent因為每次只看一個點, 如果這次的更新取到不好的點, error就會突然噴很高, 如果取到的是好的, error就會下降, 所以圖形上會有很大的起伏, 但隨著看的點越來越多, E_{in} 整體來看還是有下降的趨勢, 只是會有大型的抖動.

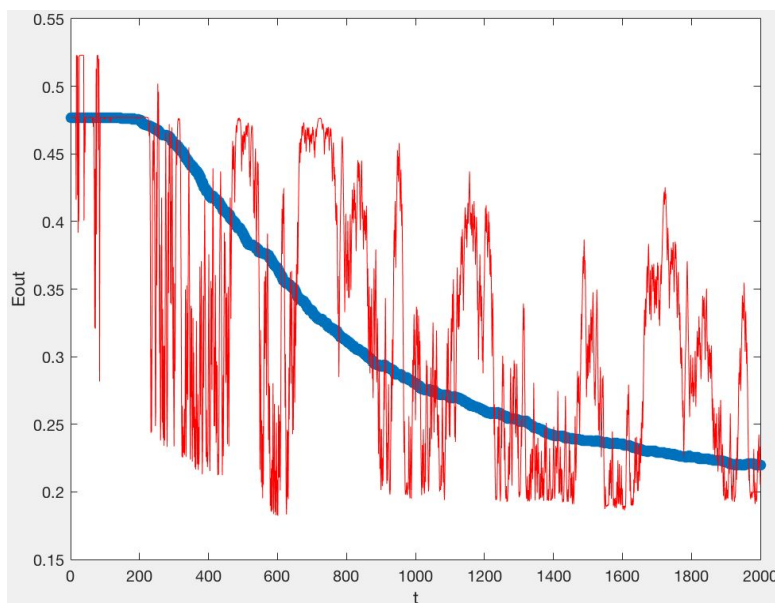
最後是以圖形來看 $\eta=0.01$ 的前200次更新大概就是 $\eta=0.001$ 全部2000次更新, 因為0.01的更新速度大概是0.001的10倍, 所以在理想的情況下(點取的沒什麼問題)大概會有同樣的趨勢。

9.

```
Eout for gradientDescent with eta = 0.001 : 0.475000  
Eout for stochasticGradient with eta = 0.001 : 0.473000  
Eout for gradientDescent with eta = 0.01 : 0.220000  
Eout for stochasticGradient with eta = 0.01 : 0.205333
```



$\eta = 0.001$



$\eta = 0.01$

以上兩張圖的藍線代表the gradient descent version, 紅線代表the stochastic gradient descent version. 這兩張圖和上一題的兩張圖有幾乎相同的現象, η 要設足夠大更新起來才有效果, 而stochastic gradient descent一樣會有大量的抖動, 但整體跑下來還是有下降的趨勢, gradient descent則是穩定下降. 而 $\eta=0.01$ 的前200次更新趨勢大概是 $\eta=0.001$

的2000次更新, 理由和上一題一樣. 另外由於這兩題得到的圖類似, 只是 E_{out} 大概都會比相對應的 E_{in} 大 $0.01 \sim 0.02$, 但差距不大, 所以應該可以說我們使用的model有學到東西.