

1.請說明你實作的 **generative model**，其訓練方式和準確率為何？

答：

照投影片的方式，分別計算 Class1 跟 class2 的 Sigma 與 mean 值，最後可以求得  $w$ ，直接與 test\_data 內積作為答案

0.842357421455115

2.請說明你實作的 **discriminative model**，其訓練方式和準確率為何？

答：train 了 500 次 iteration，每次更新  $w$  與 bias 值，即可求得最終的  $w$

0.839622641509434

3.請實作輸入特徵標準化(feature normalization)，並討論其對於你的模型準確率的影響。

答：一開始沒做 normalization 的時候，learning rate 必須條的非常小，而且準確率不一定會上升，有時候反而會 73% 27% 互相交錯，做過 normalization 之後，learning rate 可以調得比較大，而且準確度也會比較穩定的上升

4. 請實作 logistic regression 的正規化(regularization)，並討論其對於你的模型準確率的影響。

答：在 traing set 的預測準確率下降，但對於 test set，overfitting 就不會影響太多，準確率因此提高

5.請討論你認為哪個 attribute 對結果影響最大？

age,workclass,education,race,hours-per-week 感覺對於收入有很大的相關