

學號：R06942082 系級：電信碩一 姓名：黃釋平

1.請比較你實作的 generative model、logistic regression 的準確率，何者較佳？

答：

	generative model	logistic regression
public	0.84434	0.85356
private	0.84092	0.85038

2.請說明你實作的 best model，其訓練方式和準確率為何？

答：

我用 4 層神經網路，優化器選擇 Adam(lr=0.03, beta_1=0.9, beta_2=0.999, epsilon=1e-8)，訓練次數為 500，準確度為 0.85479。

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：

若沒有做 feature normalization，神經網路會跑不動，logistic 則是精準度變為 0.79766，generative model 卻神奇地變成 0.84533。

上述我只對非數位的值做 normalization，如 sex 只會是 0 和 1，就不做 normalization，原因是資料本來就位於 0~1 之間。

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：

	$\lambda = 0.1$	$\lambda = 0.01$	$\lambda = 0.001$	$\lambda = 0$
public	0.85356	0.85356	0.85343	0.85343
private	0.85014	0.85038	0.85038	0.85038

5.請討論你認為哪個 attribute 對結果影響最大？

我認為可以以 generative model 計算出的 weighted 大小，可以代表那個 feature 的重要性，其中以 fnlwgt、capital_gain 最為明顯。