學號：B05901003 系級：電機二 姓名：徐敏倩

1. 請比較你實作的generative model、logistic regression 的準確率，何者較佳？
在資料處理的部份，皆是使用已經分好的X_train, Y_train（且已特徵標準化），Logistic Regression的準確率皆較 Generative Model來的高。下表中，Generative Model是利用 Gaussian distribution的假設實作；Logistic Regression的部分，learning rate為0.01，gradient descent 的部分用adagrad，訓練 100000個 epoch。

	Training Accuracy	Testing Accuracy
Logistic Regression	0.853291	0.852645
Generative Model	0.842296	0.843860

2. 請說明你實作的best model，其訓練方式和準確率為何？
我的best model是用fully connected forward network，input data是用已經分好的X_train, Y_train，並已做特徵標準化，且切割出前三分之二作為 training set、後三分之一作為 validation set。Network中有三個 hidden layer，第一、二、三層中分別有1024, 512, 128個 neurons，而每一層都有用activation function, sigmoid，loss function為 'binary_crossentropy'，optimizer為 'adam'，且在選擇model時，是選擇在 validation set準確率最高的model。而最後在 testing data set 上的準確率為 85.88%。

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
由下表可知（以 Logistic Regression為例），當資料未經過特徵標準化的處理時，準確率不管是在 training set還是在 testing set上都較經過特徵標準化的差。除了準確率之外，在未做特徵標準化時，learning rate需要在0.0001左右，才有辦法收斂，否則便會在 trainig accuracy為0.76, 0.24之間跳，而根本不會上升；而經過特徵標準化後的資料，learning rate為0.1時，就已經可以收斂了。

		Training Accuracy	Testing Accuracy
Logistic Regression	Original	0.804275	0.806825
	Normalization	0.853291	0.852645

未經特徵標準化：learning rate為0.0001、gradient descent部分用adagrad、epoch為10000
經過特徵標準化：learning rate為0.01、gradient descent部分用adagrad、epoch為10000

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。

λ	Training Accuracy	Testing Accuracy
0.1	0.7855	0.7858
1	0.7643	0.7691
10	0.7636	0.7681
100	0.7632	0.7677

從上表中可以得知，在 λ = 0.1 的情況下，準確率較其他幾個來的稍微高一點；但在 λ = 1, 10, 100時， λ 的影響並不大，但還是可以發現，不論是 training accuracy 還是 testing accuracy，其實都會因 λ 變大而減少。

5. 請討論你認為哪個attribute 對結果影響最大？

我認為最重要的是workclass，當我去掉workclass這個attribute，以其他attribute作為訓練資料時，準確率降到80.87%左右；其次則為marital_status，當我去掉此attribute，以其他attribute作為訓練資料時，準確率降到81.75%左右。而對於其他的attribute重複相同的步驟，準確率都還是落在83%到85%之間。因此，workclass、marital_status這兩個 attribute對結果的影響最大。