

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

	private score	public score
generative model	0.73479	0.73120
logistic regression	0.85640	0.85626

logistic regression的準確率明顯優於generative model

2. 請說明你實作的best model，其訓練方式和準確率為何？

在我的best model中，我把所有continuous的參數(age, fnlwgt, capital_gain, ...)都加入一個二次項，在進行normalization 跟 regularization，最後再利用logistic regression的方式來訓練。

	private score	public score
logistic regression	0.85640	0.85626

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

	private score	public score
no normalization	0.79277	0.79729
normalization	0.85640	0.85626

使用feature normalization後可以發現模型的準確率有明顯的提昇。

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。

	private score	public score
no regularization	0.85738	0.85589
regularization	0.85640	0.85626

在我實作的模型中，有沒有使用regularization對最後的準確率影響並不大。

5. 請討論你認為哪個attribute 對結果影響最大？

為了看出哪項feature對結果的影響最大，我將各項的weight取絕對值後，前三大的值所對應到的feature分別是fnlwgt, capital gain, capital loss，因此我覺得這三項對結

果的影響最大。