

學號 B05901063 系級：電機三 姓名：黃世丞

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

	Private	Public
Generative model:	0.84117	0.84643
Logistic regression:	0.84866	0.85466

Logistic regression 的準確率較佳

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

使用原來的 **logistic regression** 再對連續的 **feature** 加上二次項，**batch\_size**=100, **epochs** 約 30000

Private: 0.85702      Public: 0.85687

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

對於某些模型(如加入二次項、**tensorflow** 疊一兩層 **hidden layer**)，沒有做 **feature normalization** 或 **batch normalization** 的話，根本 **train** 不起來，**loss** 一開始就會爆掉。但對於本來就可以訓練的模型來說，**normalization** 並沒有對準確率帶來太多影響。

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。

對於 **logistic regression**，正規化並沒有對準確率帶來太多影響，但對較複雜的模型(如兩層 **hidden layer**)，訓練過程可以明顯看到降低了 **overfit** 的情形。

5. 請討論你認為哪個 **attribute** 對結果影響最大？

把所有 **weight** 印出來之後，絕對值最大的 **feature** 是 **Capital gain**