

學號：r06942095 系級：電信碩一 姓名：劉翔瑜

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

答：

logistic regression:約為85%

generative model:約為 78%

logistic regression較佳

2.請說明你實作的best model，其訓練方式和準確率為何？

答：

使用logistic regression，將人種,婚姻狀況,所屬國家的feature消去,並concatenate‘原本’資料之平方，之後再normalize，最後將這些資料拿去train

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：

Normalize前，正確率為77%，normalize後為85%

根據gradient descent的觀念，normalize後的梯度會直指最低點，推測是因此normalize後的data會train出較好的結果

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：

Lamda=1 ---->84%

Lamda=2----->83.8%

Lamda=3----->83.4%

推測是加入lamda後，主要會影響結果的feature之weight變小，導致預測變不準確

5.請討論你認為哪個attribute對結果影響最大？

加入二次項並 normalize 後，影響最大，從原本的 77%，突破到 85%