

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 84%，突破到 85%