

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

答：

使用高斯函數當作模型，但是跑出來就是慘慘的(只有 0.84646)

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

答：

使用 gradient descent 實作 logistic regression，然後準確率大約為 0.86095 左右

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：

沒有作 feature normalization 前會連 simple baseline 都過不了，但是做了之後就能通過 simple baseline 了。

原因有可能是若沒有做 feature normalization 會使得有部分項在前期因為數字過大過小，而直接噴掉

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：

對於我的模型，logistic regression 的正規化(regularization)好像沒有很大的幫助，如果設太小，跟沒有設差不多，但是如果設太大，反而準確率會下降(0.860->0.855)

5.請討論你認為哪個 attribute 對結果影響最大？

答：

把 logistic regression 中的 w 印出來後列出絕對值前 3 大的項，發現分別是 doctorate, capital_gain, capital_gain^0.5

因此對結果影響最大的 attribute 應該是 capital_gain