學號：R05921078  系級： 電機碩1  姓名：洪立達

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

**答：**

從training data中隨機取出3萬筆data，然後比較在training data中的預測準確率。但是測試下來，準確率最高還是將training data全取，其準確率為0.8910

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

**答：**

每一次刪除一個feature後，再將剩下的feature全部拿去做training，得到越低的預測率即表示該刪除的feature越重要。利用此方法得到幾個比較重要的feature來建構我的model，其準確率為0.9018

**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：**

有用特徵標準化的準確率：0.9018

沒用特徵標準化的準確率：0.8182

因為我有使用到某些attribute的2次式來當作feature，所以不作特徵標準化的話，參數的設定上非常難調整，準確率也較低

**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：**

有用正規化的準確率：0.90181164  
沒用正規化的準確率：0.90181118

兩者沒有明顯的差異  
**5.請討論你認為哪個attribute對結果影響最大？**

利用2.所提到的方法，分析出來的結果：capital\_gain為最重要的attribute