學號：R05546022 系級： 工工所 碩二 姓名：謝立成

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

答：

Generative : 0.84475

Logistic : 0.848595

2.請說明你實作的best model，其訓練方式和準確率為何？

答：

XGBoost ensemble gradient descent boosting with each tree maximum split = 6

Accruacy: 0.87776

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：

Logistic regression: 0.85043，Accruacy increase

Generative: 0.84511，Accruacy increase

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：

Logistic regression: 0.85043，Accruacy increase

5.請討論你認為哪個attribute對結果影響最大？

fnlwgt，By using the Importance function in XGBoost, it indicates the feature in which used most frequently during node split.