學號：R06921066 系級： 電機碩一 姓名：劉宇閎

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

答：

使用logistic regression時準確率平均可以達到0.85，但是使用generative時只有0.79。

有可能是因為數據imbalance且數據亮相差過大所造成的結果

2.請說明你實作的best model，其訓練方式和準確率為何？

答：

best model 使用的是二次的w且去除一些featurec後進行min.max 標準化

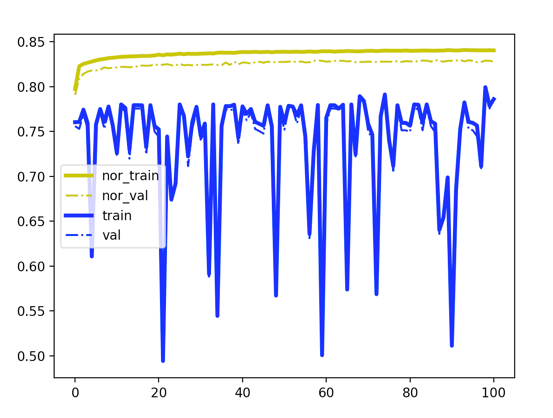
數據分割為batch size=256 並使用adam演算法來進行修正，再進行了3萬個epoch後丟上kaggle可以得到public為 private為

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：

此數據為迭代100次的結果，如果做高次迭代後亦會得到類似的圖形，

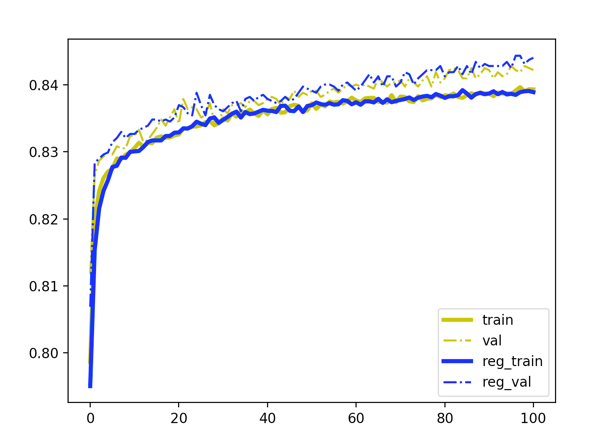
可以發現沒有做標準化時準確率會有大幅的震盪且不易收斂。



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：

使用regularization後model收斂較佳，但經過實驗後的準確率並沒有顯著的上升，推測是此種做法的模型已達全域最佳解。



5.請討論你認為哪個attribute對結果影響最大？

以結果來說capital gain影響算是最劇烈的，且在train完後的model中也可以發現他的權重植最高。