Homework 2 Report - Income Prediction

學號: r05543054 系級: 應力二 姓名: 劉禮榮

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

	private leaderboard	Public leaderboard
Generative model	0.76280	0.76474
Logistic model	0.84510 ^{win}	0.84778 ^{win}

Logistic regression 表現比較好

2. (1%) 請說明你實作的 best model, 其訓練方式和準確率為何?

實作上課講義的 Logistic regression。

- * 額外使用的 tips:
- 1.regularization
- 2.feature scaling
- 3.batch
- * 準確率

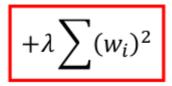
	private leaderboard	Public leaderboard
Logistic model	0.84510	0.84778

3. (1%) 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。(有關 normalization 請參考: https://goo.gl/XBM3aE)

標準化後結果比較好

	private leaderboard	Public leaderboard
raw	0.78172	0.78943
normalized	0.83429 ^{win}	0.83869 ^{win}

4. (1%) 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。(有關 regularization 請參考: https://goo.gl/SSWGhf P.35)



在 loss 後面加上這項,微分後變成

2*W,但是對結果完全沒有影響,沒變好,也沒變差

5. (1%) 請討論你認為哪個 attribute 對結果影響最大?

我把變異數太小的項都去掉,最後剩下這些

[80, 79, 78, 77, 76, 72, 68, 67, 66, 62, 61, 60, 50, 49, 47, 43, 42, 41, 37, 23, 22, 17, 10, 8, 0]

共25個,做出最好的結果,但也很難說哪個對預測的貢獻最高