

學號：R08922167 系級：資工碩一 姓名：曾民君

1. (2%) 請比較實作的 generative model 及 logistic regression 的準確率，何者較佳？請解釋為何有這種情況？

Ans: generative: 0.88241 (public), logistic regression: 0.89037, 相對來說 logistic regression 表現較好，可能原因為兩個class的資料量不平均，導致generative偏向一方。

2. (2%) 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。接著嘗試對正規項使用不同的權重 (lambda)，並討論其影響。(有關 regularization 請參考 <https://goo.gl/SSWGhf> p.35)

Ans: 固定條件: learning rate: 0.2, epochs: 20, batch_size: 8, validation set 20%

lambda: 0 accuracy: 0.88981

lambda: 0.001 accuracy: 0.88795

lambda: 0.0005 accuracy: 0.88963

這邊看起來加入 regularization 表現沒有比較好，可能是原本沒有overfitting的問題，所以加入後沒有太大成效

3. (1%) 請說明你實作的 best model，其訓練方式和準確率為何？

Ans: 使用 logistic regression, learning rate: 0.18, epochs: 80, batch_size: 8

validation set 佔 20%，只對特徵非二元的做normalization，並在訓練期間紀錄在 validation set 表現最好的 weight，此方法得到的準確率為 0.89110

4. (1%) 請實作輸入特徵標準化 (feature normalization)，並比較是否應用此技巧，會對於你的模型有何影響。

Ans: 沒有特徵標準化的情況整體準確率會下降 10 ~ 12%，同時表現較好的權重會高達 50 以上，而有特徵標準化的情況表現較好的權重大約只會落在 1.26左右