

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

答：

$$z = \frac{(\mu^1 - \mu^2)^T \Sigma^{-1} x}{w^T} - \frac{\frac{1}{2}(\mu^1)^T \Sigma^{-1} \mu^1 + \frac{1}{2}(\mu^2)^T \Sigma^{-1} \mu^2 + \ln \frac{N_1}{N_2}}{b}$$

$$P(C_1|x) = \sigma(w \cdot x + b)$$

先用訓練資料求出  $w^T$  跟  $b$ ，再用  $w^T$  跟  $b$  帶入下面的公式求出預測的  $P$ ，上傳 kaggle 的分數為 0.84128。

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

答：

$$w_i \leftarrow w_i - \eta \sum_n - \left( \hat{y}^n - \underline{f_{w,b}(x^n)} \right) x_i^n$$

與作業一的作法相同， $y$  帶入 0 或 1，在  $\lambda$  為 0 且 iteration 為 3000 時結果最好

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：

在連續的 attribute 上作 normalization，對準確率有提升

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：

5.請討論你認為哪個 **attribute** 對結果影響最大？