

Homework 2 Report

資工所碩一 R07922120 陳禹達

Problem 1. (1%) 請簡單描述你實作之 logistic regression 以及 generative model 於此 task 的表現,試著討論可能原因。

A:

當 training data 數量夠多時, logistic regression 的表現會較佳, 而數量少時, Generative model 的表現則較佳, 因為 generative model 的表現會類似補足其他 Data, 因此在數量少時可以較 logistic regression 較佳, 而數量夠多時, logistic regression 則會反超過 generative model。

	Data:100	Data:20000
Logistic regression	75.61%	82.05%
Generative model	77.37%	81.77%

Problem 2. (1%) 請試著將 input feature 中的 gender, education, marital status 等改 **F** one-hot encoding 進行 training process,比較其模型準確率及其可能影響原因。

A:

	Training accuracy	Testing accuracy
One-hot-encoding	82.01%	82.05%
NO One-hot-encoding	81.93%	82.00%

將 Input feature 改成 one-hot-encoding 的 accuracy 較佳, 分析其可能原因為, 在資料中, gender, education, marital status 在不同的類別中, 並沒有大小之分, 也無連續關係, 因此不能用數字去呈現, 所以需要將每一個類別提出來判斷為 1 或是 0 (如: gender:1,2 → male:0,1 and female:0,1), 才是合理的作法。

Problem 3. (1%) 請試著討論哪些 input features 的影響較大(實驗方法不限)。

A:

我所實驗的方法為, 先利用全部的 Feature 做出一個 model, 得到 training accuracy, 接著對每個 feature 分別做刪減做出另一個 Model, 得到另一個 training accuracy, 並與用全部 feature 的 accuracy 做比較, 若與使用全部 feature 的 accuracy 差距大於 0.1%則判斷為影響較大的 feature。

表格中: 第一列:刪除的 feature 第二列:Training accuracy

第三列: (Training accuracy-Total_feature_accuracy)/Total_feature_accuracy

None	LIMIT_BAL	SEX	EDUCATION	MARRIAGE	AGE
82.01%	81.91%	81.95%	82.01%	0.8198	0.8195
0%	-0.122%	-0.073%	0%	-0.037%	-0.073%

PAY_0	PAY_2	PAY_3	PAY_4	PAY_5	PAY_6
80.47%	81.97%	81.99%	81.92%	81.93%	81.85%
-1.878%	-0.049%	-0.024%	-0.11%	-0.098%	-0.195%

BILL_AMT1	BILL_AMT2	BILL_AMT3	BILL_AMT4	BILL_AMT5	BILL_AMT6
81.94%	81.97%	81.93%	81.96%	82.00%	82.00%

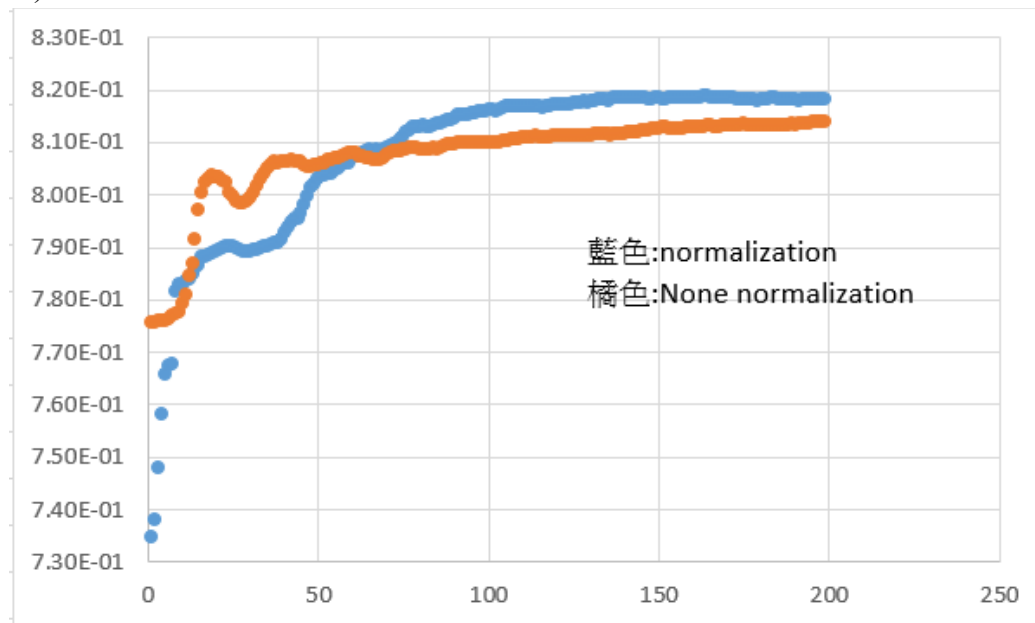
-0.085%	-0.049%	-0.098%	-0.073%	-0.012%	-0.012%
PAY_AMT1	PAY_AMT2	PAY_AMT3	PAY_AMT4	PAY_AMT5	PAY_AMT6
81.96%	81.97%	81.88%	81.89%	81.88%	82.00%
-0.06%	-0.049%	-0.159%	-0.146%	-0.159%	-0.012%

最終所得到的影響較大的 Feature 為 LIMIT_BAL, PAY_0, PAY_6, PAY_AMT3, PAY_AMT4, PAY_AMT5。

Problem 4. (1%) 請實作特徵標準化 (feature normalization),討論其對於你的模型準確率的影響。

A:

在最終結果中，不管有沒有做 normalization 的 Testing accuracy 是相差不大的，因此我認為有無做 normalization 的影響並不大，但如果有做 normalization 的話，training 的速度會較快，也就是 logistic regression 會較快達到 minimum，可以參考下圖，我們取前兩百個 iteration 做比較，可以看出有做 normalization 在 training 的過程中較快達到 minimum(accuracy 較高)。



Problem 5. (1%)

$$5. f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$

$$\text{let } u = \frac{x-\mu}{\sigma}$$

$$\Rightarrow du = \frac{dx}{\sigma}, dx = \sigma du$$

$$\text{let } Q = \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \sigma du = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

$$\Rightarrow Q^2 = \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \right) \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{(u^2+v^2)}{2}} du dv$$

$$\text{let } u = r \cos \theta, v = r \sin \theta, 0 \leq \theta \leq 2\pi, 0 \leq r < \infty$$

$$\Rightarrow Q^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-\frac{(r^2 \cos^2 \theta + r^2 \sin^2 \theta)}{2}} |J| dr d\theta, |J| = \begin{vmatrix} \frac{du}{dr} \frac{du}{d\theta} & \frac{du}{dr} \frac{dv}{d\theta} \\ \frac{dv}{dr} \frac{du}{d\theta} & \frac{dv}{dr} \frac{dv}{d\theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \cos^2 \theta + r \sin^2 \theta$$

$$\Rightarrow Q^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-\frac{r^2}{2}} r (\cos^2 \theta + \sin^2 \theta) dr d\theta$$

$$\because \cos^2 \theta + \sin^2 \theta = 1$$

$$\therefore Q^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-\frac{r^2}{2}} r dr d\theta$$

$$\text{let } w = r^2$$

$$\Rightarrow dw = 2r dr$$

$$\Rightarrow r dr = \frac{dw}{2}$$

$$\Rightarrow Q^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-\frac{w}{2}} \frac{dw}{2} d\theta$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} \left(-2e^{-\frac{w}{2}} \right) \Big|_0^{\infty} d\theta$$

$$= \frac{1}{2\pi} \int_0^{2\pi} (0 - (-1)) d\theta$$

$$= \frac{1}{2\pi} \int_0^{2\pi} 1 d\theta$$

$$= \frac{1}{2\pi} (\theta \Big|_0^{2\pi})$$

$$= \frac{1}{2\pi} (2\pi - 0)$$

$$= 1$$

$$Q^2 = 1 \Rightarrow Q = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1, \text{得證} \#$$

Problem 6. (1%)

$$6. (a) \frac{\partial E}{\partial z_k} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial z_k} = g'(z) \frac{\partial E}{\partial y_k} \quad \#$$

$$(b) \frac{\partial E}{\partial z_j} = \sum_k \frac{\partial E}{\partial z_k} \frac{\partial z_k}{\partial z_j}$$

$$\because z_k = \sum_j w_{jk} y_j = \sum_j w_{jk} g(z_j)$$

$$\therefore \frac{\partial z_k}{\partial z_j} = w_{jk} g'(z_j)$$

$$\Rightarrow \frac{\partial E}{\partial z_j} = \sum_k w_{jk} g'(z_j) \frac{\partial E}{\partial z_k} \quad \#$$

$$(c) \text{ let } f = \frac{\partial E}{\partial z_j}$$

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial z_j} \frac{\partial z_j}{\partial w_{ij}} = \frac{\partial E}{\partial z_j} \frac{\partial \sum_k w_{kj} y_k}{\partial w_{ij}} = y_i \frac{\partial E}{\partial z_j} \quad \#$$