Homework 2 Report

學號: R07946007　　系級:資料科學碩一　　姓名:陳庭安

1. (1%) 請說明你實作之 logistic regression 以及 generative model 於此 task 的表現,並試著討論造成此差異及可能原因。

Model	Training Accuracy	Validating Accuracy	Testing Accuracy (Public / Private)
Logistic	0.77775	0.83	0.78060 / 0.78120
Generative	(much runtime)	0.784107875	0.78060 / 0.78140

Generative model 在 validation set 的 accuracy 雖然比較低，但在 testing 和 Logistic model 有一樣的表現。Logistic model 在此稍微高估了分類預測表現，Generative model 則無高估的問題，因為 Generative model 在此是由常態分佈生成機率值預測，較不會受離群值影響。

2. (1%) 請試著將 input feature 中的 gender, education, marital status 等改為 one-hot encoding 進行 training ,並比較其與原本的差異及其可能原因。

Model (One-hot)	Training Accuracy	Validating Accuracy	Testing Accuracy (Public / Private)
Logistic	0.772	0.78475	0.78060 / 0.78120
Generative	(much runtime)	0.784107875	0.78060 / 0.78140

類別變數經過 One-hot encoding 後，Generative model 的 accuracy 無明顯變動；然而 Logistic model 的 validating accuracy 下修許多，改善了高估的問題。因若將類別變數視作連續型變數處理，各類別的數值順序大小將影響預測結果。

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

Logistic model (data scaled)

	Most influential (top 3)			Least influential		
Features	PAY_AMT1	PAY_AMT2	PAY_AMT3	BILL_AMT1	PAY_AMT6	PAY_AMT5

Weights	-54.65	-48.26	-30.48	-0.13	0.33	1.10

Delay 後 1~3 個月的 Payment 影響最大，值越大，表示有付款，較不會是不還款的用戶；反之，3 個月內還沒有甚麼 Payment，很可能就根本不會還了。影響預測結果較不顯著的除了 9 月份用戶的 Bill payment amount (BILL_AMT1)，還有 Delay 後 5~6 個月的 Payment，兩類用戶這些數值分布差不多。

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

Model (Scale)	Training Accuracy	Validating Accuracy	Testing Accuracy (Public / Private)
Logistic	0.7276	0.81	0.73540 / 0.73280
Generative	(much runtime)	0.78425	0.78060 / 0.78120

Model (One-hot+ Scale)	Training Accuracy	Validating Accuracy	Testing Accuracy (Public / Private)
Logistic	0.746	0.75	0.74435 / 0.74640
Generative	(much runtime)	0.784107875	0.78060 / 0.78140

以平均數、標準差做平移、伸縮的標準化後，Logistic model 預測結果會比較差，因為在沒考慮 outlier 的情況下直接標準化，會有明顯 bias；Generative model 表現沒有較差，可能是因為是由常態分配生成機率值，結果較穩定。

5.

令 $I = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx$

$\Rightarrow I^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \cdot \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dx\,dy$

$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \Rightarrow J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$

$\Rightarrow I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}}\, r\,dr\,d\theta = \int_0^{2\pi}\int_0^{\infty} \frac{1}{2} e^{-\frac{r^2}{2}} d(r^2)\,d\theta$

$= \int_0^{2\pi} -e^{-\frac{r^2}{2}} \Big|_0^{\infty} d\theta = \int_0^{2\pi} 1\,d\theta = 2\pi$

$\Rightarrow I = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$

$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \overset{令 z = \frac{x-\mu}{\sigma}}{=\!=} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma \cdot dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz$

$= \frac{1}{\sqrt{2\pi}} \cdot I = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1$. 得證.

6.

(a) $$\frac{\partial E}{\partial z_k} = \frac{\partial E}{\partial g(z_k)} \frac{\partial g(z_k)}{\partial z_k} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial z_k}.$$

(b) $$\frac{\partial E}{\partial z_j} = \frac{\partial E}{\partial g(z_k)} \cdot \frac{\partial g(z_k)}{\partial z_k} \cdot \frac{\partial z_k}{\partial y_j} \cdot \frac{\partial y_j}{\partial z_j}$$

$$= \sum_k \left(\frac{\partial E}{\partial g(z_k)} \frac{\partial g(z_k)}{\partial z_k} \cdot w_{jk} \right) \cdot \frac{\partial g(z_j)}{\partial z_j} = \left(\sum_k \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial z_k} w_{jk} \right) \frac{\partial y_j}{\partial z_j}$$

(c) $$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial z_j} \frac{\partial z_j}{\partial w_{ij}} = \left(\sum_k \frac{\partial E}{\partial g(z_k)} \frac{\partial g(z_k)}{\partial z_k} \cdot w_{jk} \right) \frac{\partial g(z_j)}{\partial z_j} \cdot g(z_i)$$

$$= \left(\sum_k \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial z_k} w_{jk} \right) \frac{\partial y_j}{\partial z_j} \cdot y_i$$