Homework 2 - Credit Card Default Payment Prediction Report Problem Set

學號： B05901082 系級： 電機三 姓名: 楊晟甫

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

	public_score	private_score
generative	81.64	81.50
logistic	82.02	82.14

相較於generative model，使用one-hot encoding以及normalization之後的logistic regression 的 performance 似乎比較好，其關鍵可能是在one-hot encoding，因為有一些離散的label我們不能假設他在數值上有大小的關係，像是gender, education, martial status, 所以使用one-hot encoding後會好很多，可以從第二題看到沒有使用one-hot的logistic的performance是比generative還差的。

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

	public_score	private_score
without one-hot	80.06	80.00
with one-hot	82.02	82.14

如同第一題提及，我們不能假設某些特定的feature值會跟我們的output有更大的關聯，像是性別、教育水平以及婚姻狀態，假設男生/女生，大學以上/大學以下等會有較好的信用水準，以常理來推斷是不合理的。

Problem 3. (1%) 請試著討論哪些 **input features** 的影響較大**(**實驗方法沒有特別限制**,**但

請簡單闡述實驗方法)。

	without feature	with feature
age	82.02	81.98
LIMIT_BAL	82.02	82.02
SEX	82.10	82.02
EDUCATION	82	82.02
PAY_	77.78	82.02
MARRIAGE	81.96	82.02

以一次抽掉一個feature的方式去看public_score的變化，可以發現只有拔掉SEX之後的performance比較好，從現實層面考量的確是合理的因素，性別應不至於影響一個人的信用水準，而PAY_是重要的feature，若是拔掉performance會很差，看起來也是合理的，因為其代表一個人的付款狀況。

	without augmentation	with augmentation
BILL_AMT	82.10	82.08
PAY_AMT	82.10	82.10

我把bill_amt跟pay_amt這兩個feature做平方之後當作新的feature，沒有抽掉的原因是這兩項自然是會影響model的，然而擴增data之後似乎對performance沒有影響，還有下降的趨勢。

Problem 4. (1%) 請實作特徵標準化 (feature normalization),並討論其對於模型準確率

的影響與可能原因。

	public_score	private_score
without normalization	77.75	77.72
with normalization	82.02	82.14

有無normalization的performance其實差蠻多的，原因是有幾個feature數值特別的大，加上有做one-hot encoding，如果沒有normalization的話會讓weight dot x 之後的數值過大，sigmoid之後幾乎都在1，讓model嚴重失真。

Problem 5.6. (2%)

No:

Date: /

HW2 電機三 B05901082 楊晟甫

5、 $\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx\right)^2 = \frac{1}{2\pi\sigma^2}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \int_{-\infty}^{\infty} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \rightarrow u=\frac{x-\mu}{\sqrt{2}\sigma}\ \sqrt{2}\sigma du = dx$

$= \frac{1}{2\pi\sigma^2}\cdot 2\sigma^2 \int_{-\infty}^{\infty} e^{-u^2} du \int_{-\infty}^{\infty} e^{-v^2} dv = \frac{1}{\pi}\iint_{-\infty}^{\infty} e^{-(u^2+v^2)} du dv = \frac{1}{\pi}\int_0^{2\pi}\int_{-\infty}^{\infty} e^{-r^2} r dr d\theta$

$= \frac{1}{\pi}\cdot 2\pi \cdot \frac{1}{2}\cdot \int_0^{\infty} e^{-t} dt = \pi\cdot(e^0 - e^{-\infty}) = 1$

$\therefore \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma}} dx = 1$ (for all probability ≥ 0, integral $\neq -1$)

6、 (a) $\frac{\partial E}{\partial z_k} = \frac{\partial E}{\partial y_k}\cdot\frac{\partial y_k}{\partial z_k} = \frac{\partial E}{\partial y_k}\cdot\frac{\partial g(z_k)}{\partial z_k}$ #

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

(c) $\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_k}\cdot\frac{\partial y_k}{\partial z_k}\cdot\frac{\partial z_k}{\partial y_j}\cdot\frac{\partial y_j}{\partial z_j}\cdot\frac{\partial z_j}{\partial w_{ij}}$ #