Homework 2 Report

Professor Pei-Yuan Wu EE5184 - Machine Learning

學號：R07922162 系級：資工碩一 姓名：胡嘉祐

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。

	Kaggle score	loss
logistic regression	0.81820	0.44994
generative model	0.82100	0.44777

實作的feat皆一樣，使用原本的feature加上 limit_ball, pay_0, bill_amt, pay_amt的二次方，以及limit_ball, pay_0, bill_amt_0的三次方，並對所有data做normalization。

Logistic regression: 初始w,b=0,learning rate =0.5 with adagrad,

直接使用20000筆data train,每100筆更新一次w,b 共做1000次iteration。

Generative model: 標準的slide提供的方法，並沒有做額外的延伸。

由此可見generative model 得到較好的結果，有可能是我的logistic regression寫得不夠完善，也有可能是data本身有noise，label並不正確，而generative model能自動忽視掉這個部份。

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

做one-hot encoding，其結果有好一點點而已，因為後來有進行normalization，所以使one hot的效果並不顯著

	Kaggle score	loss	Kaggle score	loss
One hot	yes	yes	no	no
logistic regression	0.81140	0.45103	0.81800	0.44538
generative model	0.79720	0.47336	0.81580	0.467629

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

使用generative model進行實驗，每次刪掉一個feature，之後進行標準化。以下表格取score下降較多的model。

發現pay與pay_amt1下降較多，其影響也較大。

	Kaggle score	loss
normal	0.81580	0.46762
Pay_0	0.79560	0.46581
Pay_2	0.79040	0.46712
Pay_3	0.80020	0.46603
Pay_4	0.80040	0.46772
Pay_5	0.80540	0.46749
Pay_6	0.80120	0.46668
PAY_AMT1	0.81380	0.46774

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。

使用特徵標準化對於training的結果，會有非常顯著的影響

有些feature像是limit_ball, bill_amt, pay_amt數值較其他數值而言大非常多，因此如果不做標準化，則會使準確度下降非常多，我的logistic的model，甚至無法training。

僅使用23個feature

	Kaggle score	loss	Kaggle score	loss
Feature Normalization	yes	yes	no	no
Logistic Regression	0.82140	0.44538	0.2194	21.49004
generative model	0.81580	0.46762	0.78120	0.68878

Problem 5. (1%) （在下面兩頁）

5.

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

∵ by chain rule

(b)

$$\frac{\partial E}{\partial z_j} = \frac{\partial y_j}{\partial z_j}\frac{\partial E}{\partial y_j} = g'(z_j)\cdot\left[\frac{\partial z_1}{\partial y_j}\frac{\partial E}{\partial z_1} + \frac{\partial z_2}{\partial y_j}\frac{\partial E}{\partial z_2} \cdots + \frac{\partial z_k}{\partial y_j}\frac{\partial E}{\partial z_k}\right]$$

∵ by chain rule

$$= g'(z_j)\cdot\left[w_{j1}\frac{\partial E}{\partial z_1} + w_{j2}\frac{\partial E}{\partial z_2} + \cdots + w_{jk}\frac{\partial E}{\partial z_k}\right]$$

∵ $\frac{\partial y_j}{\partial z_j} = g'(z_j)$

$$= g'(z_j)\sum_{i=1}^{k} w_{ji}\frac{\partial E}{\partial z_i}$$ #

(c) ∵ by chain rule

$$\frac{\partial E}{\partial w_{ij}} = \underbrace{\frac{\partial z_j}{\partial w_{ij}}}_{y_i}\cdot\underbrace{\frac{\partial E}{\partial z_j}}_{\text{from (b)}} = y_i\cdot g'(z_j)\sum_{i=1}^{k} w_{ji}\frac{\partial E}{\partial z_i}$$ #

Problem 6. (1%)

$f(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ 設 $u = \frac{x-\mu}{\sigma}$, $du = \frac{dx}{\sigma}$

, $dx = \sigma du$

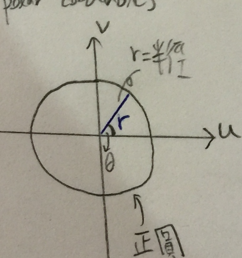
$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\cancel{\sigma}} e^{-\frac{1}{2}u^2} \cancel{\sigma} du$

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$

設 $v=u$

$f(x)$取平方： $f^2(x) = \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du\right)\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2} dv\right)$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(u^2+v^2)} du dv$ ∵ by polar coordinates



$f^2(x) = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}(r^2\cos^2\theta + r^2\sin^2\theta)} |J| dr d\theta$

$|J| = \begin{vmatrix} \frac{\partial u}{\partial \theta} & \frac{\partial u}{\partial r} \\ \frac{\partial v}{\partial \theta} & \frac{\partial v}{\partial r} \end{vmatrix} = \begin{vmatrix} -r\sin\theta & \cos\theta \\ r\cos\theta & \sin\theta \end{vmatrix} = |-r\sin^2\theta - r\cos^2\theta|$ $0 \le \theta \le 2\pi$, $0 \le r \le \infty$

∵ $\sin^2\theta + \cos^2\theta = 1$ $u = r\cos\theta$, $v = r\sin\theta$

$= r$

$f^2(x) = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}r^2} r dr d\theta$, 設 $w = r^2$, $dw = 2r dr$, $r dr = \frac{dw}{2}$

$= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}w} \frac{1}{2} dw d\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{2}\left[-2e^{-\frac{1}{2}w}\Big|_0^{\infty}\right] d\theta$ $[0-1] = -1$

$= \frac{1}{2\pi}\int_0^{2\pi} d\theta = \frac{1}{2\pi}\theta\Big|_0^{2\pi} = \frac{1}{2\pi}[2\pi - 0] = 1$ $f^2(x) = 1$

∵ integral.

$f(x) = 1$ 得證 #.