HW2 Report

B05901009 電機三 高瑋聰

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。

本次作業我實作了最基本的 logistic regression (1/(1+e^-z), z= $w^T x+b$)、generative model (假設每一維分佈皆為 Gaussion)，以及增加一層 hidden layer 的 logistic regression。這三個 model 中，有一層 hidden 的 model 較為複雜，performance 也較好，但容易在 training set 上 overfit，而 generative 的表現最差，可能是預設為 gaussian 的前提限制太大，因此無法獲得太好的效果。

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。

若將 gender, education 等參數做 one-hot encoding，model 的準確率會比沒做 one-hot 十多約 1%，可見利用 one-hot 可以讓不連續的類別變量變得更加合理，也能夠從中獲得更多有用的資訊 (原本用整數的表達方式會讓不同類別有大小關係，因此不合理)。

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

我在 training 時若將類別變量 (例如 sex、education 等)移除，若只去掉 sex 或 education 的話 model 的結果與加入這些 feature 時差不多，但若將 PAY_0~PAY_6 也都去掉，則 model 的表現就掉至 baseline(0.78)，反之我將連續變量 (PAY_AMT 等) 加上二次的 feature 與三次的 feature，model 的準確率都沒有上升。可見在這份 data 中，PAY_0~PAY_6 等的影響較 PAY_AMT、BILL_AMT 等變量來得重要。

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。

特徵標準化我認為是這次作業中影響 logistic model 最大的部份，在有標準化的情況下，大約能有 81~82%的準確率，但是一旦拿掉特徵標準化，training loss 會直接變成 nan (我是使用 Adam 當 optimizer)，準確率也會只剩下 70%，可見特徵標準化對於 gradient 與 loss 的穩定與最後準確率的表現有十分深遠的影響。

Problem 5. 6. 請見下頁圖片

B05901009 電機三 高瑋聰

5、令 $f(x,y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^{\infty}\int_0^{2\pi} \frac{1}{2\pi\sigma^2} e^{-\frac{(r\cos\theta)^2+(r\sin\theta)^2}{2\sigma^2}} r\,d\theta\,dr \quad \begin{pmatrix} x=\mu+r\cos\theta \\ y=\mu+r\sin\theta \end{pmatrix}$$

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

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

$$\Rightarrow \int_{-\infty}^{\infty} f(x)dx = 1 \quad ※ \quad (\because f(x) \geq 0 \ \forall -\infty < x < \infty)$$

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

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

(c) $\frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial z_j} \cdot \frac{\partial z_j}{\partial W_{ij}} = \frac{\partial E}{\partial y_k} \cdot g'(z_k) \cdot W_{jk} \cdot g'(z_j) \cdot \frac{\partial z_j}{\partial W_{ij}}$ ※