Homework 2 Report Problem Set

Professor Pei-Yuan Wu

EE5184 – Machine Learning

電子所碩一 R07943123 馬咏治

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

Ans:

Feature Select: LIMIT_BAL, PAY_0

	Score in Kaggle
Logistics Regression	0.8200
Generative Model	0.8200

首先我們對這兩個 feature 進行 normalization，然後丟進去 Logistics Regression 進行 training。

發現 Generative Model 和 logistics regression 這兩種方法得出的 result 竟然完全一樣。

我想這兩種方法得出一樣的結果可能是因為這兩種方法找到的 optimal point 非常接近，所以我們在切、分類的結果才會完全一樣

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

Feature Select: ALL

	Score in Kaggle
w/ one hot encoding	0.78120
w/o one hot encoding	0.77640

Feature Select: PAY_0

	Score in Kaggle
w/ one hot encoding	0.82100
w/o one hot encoding	0.82000

從上面兩個 table 看到結果不太一致，我想當選對 feature 時，one hot encoding 會幫助我們分類（從第二張圖可以看到），但選到很多沒用的 feature 時，就比較看不出效果。

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）

Feature ：ALL

實驗方式：把所有 feature 丟進去 logistic regression 裡 train，train 完後觀察每個 feature 的 weight

FEATURE	ABS(WEIGHT)
PAY_0	80.9493085
PAY_2	17.6640955
EDUCATION	15.8737184
MARRIAGE	7.47136739
PAY_3	6.5063611
PAY_5	6.33440712
AGE	1.94342393
SEX	1.5006359
PAY_4	1.26719254
PAY_6	0.575216753
PAY_AMT2	0.005693991
PAY_AMT3	0.004656391
PAY_AMT1	0.004349175
PAY_AMT5	0.003901394
BILL_AMT1	0.003066846
PAY_AMT4	0.002418185
LIMIT_BAL	0.00175016
BILL_AMT4	0.001047222
PAY_AMT6	0.000800981
BILL_AMT2	0.000693118
BILL_AMT5	0.000645673
BILL_AMT6	0.000641515
BILL_AMT3	0.000530806

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。

Feature: ALL

	Score in Kaggle
w/ Normalization	0.78120
w/o Normalization	0.79440

w/o normalization 的結果比預期的來的好，但是在 training 的過程發生了 overflow 的問題，我想原因在於有些 feature 的值太大，造成 overflow。

Problem 5. (1%) The Normal (or Gaussian) Distribution is a very common continuous probability distribution. Given the PDF of such distribution

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

please show that such integral over $(-\infty, \infty)$ is equal to 1.

Since the Gaussian integral

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}$$

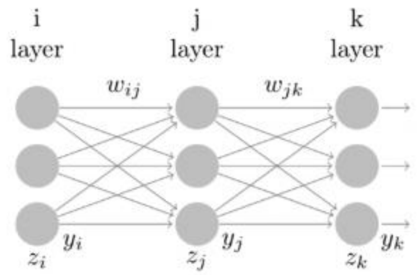
So,

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

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} \cdot e^{-\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)^2} \; d\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)$$

$$= \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1_{\#}$$

Problem 6. (1%) Given a three layers neural network, each layer labeled by its respective index variable. I.e. the letter of the index indicates which layer the symbol corresponds to.



For convenience, we may consider only one training example and ignore the bias term.Forward propagation of the input z_i is done as follows. Where $g(z)$ is some differentiable function (e.g. the logistic function).

$$y_i = g(z_i)$$
$$z_j = \sum_i w_{ij} y_i$$
$$y_j = g(z_j)$$
$$z_k = \sum_j w_{jk} y_j$$
$$y_k = g(z_k)$$

Derive the general expressions for the following partial derivatives of an error function E, also sime differentiable function, in the feed-forward neural network depicted. In other words, you should derive these partial derivatives into "computable derivative" (e.g. $\frac{\partial E}{\partial y_k}$ or $\frac{\partial z_k}{\partial w_{jk}}$).

$$(a)\frac{\partial E}{\partial z_k} \quad (b)\frac{\partial E}{\partial z_j} \quad (c)\frac{\partial E}{\partial w_{ij}}$$

a.) Let $\mathrm{E} = \mathrm{f}(y_k)$

$$\frac{\partial \mathrm{E}}{\partial z_k} = \frac{\partial \mathrm{E}}{\partial \mathrm{f}} \cdot \frac{\partial \mathrm{f}}{\partial y_k} \cdot g'(z_k)$$

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

c.) $$\frac{\partial \mathrm{E}}{\partial w_{ij}} = \frac{\partial \mathrm{E}}{\partial z_j} \cdot \sum_i y_i$$