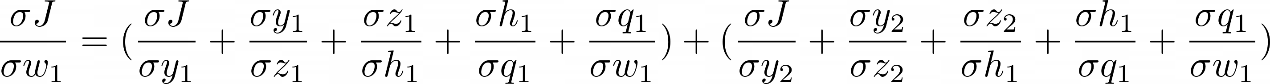
Final Exam  
Student ID: 109598001  
Student Name: 鄭濠營

1. P1 = (1, 2), C = +1; P2 = (2, 1.5), C＝+1; P3 = (1.5, 1) ,C = +1; P4 = (1, 1), C = -1; P5 = (2, 1), C = -1, 假設initial weight= 0.2, 0.2, 0.2, 0.2, 0.2。考慮可能的threshold去做分類判斷，1-feature threshold = 1.2, 1.8, 2-feacture threshold = 1.2, 1.8。之後設定簡單分類判斷: 如果feature value < threshold 分類為正，反之為負。  
   1st iteration:  
   (1) 1-feacture classification is threshold 1.2, and P1(+), P2(-), P3(-), P4(+), P5(-)。  
   對應的正確與否為: P1(T), P2(F), P3(F), P4(F), P5(T)  
   (2) 2-feacture classification is threshold 1.2, and P1(-), P2(-), P3(+), P4(+), P5(+)。  
   對應的正確與否為: P1(F), P2(F), P3(T), P4(F), P5(F)  
   (3) 1-feacture classification is threshold 1.8, and P1(+), P2(-), P3(+), P4(+), P5(-)。  
   對應的正確與否為: P1(T), P2(F), P3(T), P4(F), P5(T)  
   (4) 2-feacture classification is threshold 1.8, and P1(-), P2(+), P3(+), P4(+), P5(+)。  
   對應的正確與否為: P1(F), P2(T), P3(T), P4(F), P5(F)  
   可選擇 (3) 或 (4) 以Error最小的作為更新權重的參數值。Weight Error為0.2\*2 = 0.3。這裡選(3)。  
   選用平均作為Update Weight的手段:  
   (0.2/3, 0.2/2, 0.2/3, 0.2/3, 0.2/2) = (0.06…, 0.1, 0.06…, 0.06…, 0.1)  
   Decision Tree:  
   Threshold < 1.8  
    |

|------------  
 | |  
classify:+(正) classify: -(負)

1. PCA是降維的分析方法，將高維度的資料投射到低維度的空間。題目提及2-D的均勻分布資料，其會投射1-D的維度，也就是一條線上。其eigenvector的方向會縮限在一條線上，eigenvalue的值皆是統一長度，原由是均勻分布，落在一條線上也是呈現等距的情形。
2. (1) output = wx+b, bias = 0. q1 = x1w1+x2w2=1; q2=x1w3+x2w4=0, activation function is ReLU, so h1 = 1, h2 = 0。 z1 = h1w5+h2w6 = 1, z2 = h1w8+h2w7 = 1。y=z, so forward computation is y1 = z1 = 1, y2 = z2 = 1。  
   (2) Backward computation  
   (2\*(1/2)\*y1\*w5\*(1 0)\*x1)+(2\*(1/2)\*y2\*w8\*(1 0)\*x1) = (2\*(1/2)\*1\*1\*1\*1)+(2\*(1/2)\*1\*1\*1\*1) = 2,  2\*0.1(learning rate) = 2。
3. 1023\*61\*3\*3\*16 = 8,986,032 (conv) 🡪 1021\*59\*3\*3\*16\*16=138,790,656 (conv) 🡪 no connection (max pool) 🡪 253\*12\*3\*3\*16\*16 = 6,994,944 🡪 no connection (max pool) 🡪 16\*63\*3\*128 = 387,072 🡪 128\*2 = 256  
   conv = image\_size\*filter\_size\*channel\*filter\_number\*previous\_filter\_number  
   Fully connection = input\_num\*hidden\_num & hidden\_num\*output\_num
4. The validation error is high, so that means high bias (overfitting)。  
   (1) 我會選擇增加hidden的神經元數量來增加模型的複雜度，進而試著降低bias。  
   (2) 選擇添增或減少dropout layer的考量在，添增dropout我會考量的點在於是否模型計算的參數過於龐大，造成計算負荷，或模型需要減少複雜度的情形下；減少dropout，則是考量參數過少，不適再繼續減少，又或者著已趨近balance則做微調，或模型需要提升複雜度的情形。  
   (3) 我認為不會，cross-entropy在於尋找資訊量最豐富的一個目的，利用比較圖像的資料量豐富與否來作為其分類依據，相當適合圖像問題 (因多個像素，而哪個圖的像素最能取得豐富的資訊來判斷)；Mean square error (MSE) 是以距離式的方式去驗證離正確答案差距多少。本題是圖像分類問題，我認為並不適合使用MSE作為cost function，MSE是常見的坐標系方式以距離差距做評估要素，但顯然圖像問題難以能呈現或轉為坐標系的問題。
5. (1) 需要訓練normal的model以及fraud的model。總計兩個model。  
   (2) normal的題目提到總共1000筆資料，其資料是正常的credit card；fraud總共100筆資料，其資料是假的credit card。  
   (3) 我會以training 70%, evaluation 20%, and test 10%的比例去做訓練。  
   (4) 使用Mean Square Error (MSE)作為cost function，以模型訓練出的最終結果accuracy用+-5的間距作為lower bound (threshold)來判斷是normal or fraud的credit card。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Threshold | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
| TPR | 1 | 1 | 1 | 1-(1\*(1/6))/2 約 0.91 | 1-(1\*(1/3))/2 約0.83 | 1-(0.5\*4)/2/2 約 0.5 | 0 |
| FPR | 1 | 0.25\*3= 0.75 | 0.25\*2= 0.5 | 0.25\*1= 0.25 | 0 | 0 | 0 |

TPR = TP/TP+FN  
Code N 是均勻機率分布，pdf- 1/(b-a): 1/1-(-3) = 0.25，針對FPR  
Code P 是三角形分布: pdf- c(peak point): 2/(b-a): 2/3-(-1)=0.5，f(x) = 2(x-a)/((b-a)(c-a)), a<=x<=c; f(x) = 2(b-x)/((b-a)(b-c)) c<=x<=b針對TPR