Homework 2

108B (5403) Machine Learning　王傳鈞　0416047

本次作業所使用到的程式碼，都已經上傳到[GitHub](https://github.com/a2468834/Machine_Learning_Grad/tree/master/HW2/MATLAB%20code)當中，若有需要參考其詳細內容，歡迎點擊連結直接前往網頁瀏覽。例如：「Q1\_plot.m」的MATLAB source code代表是第一題所使用到的繪圖程式碼，「Q2\_a.m」的MATLAB code則代表第二題的(a)小題所使用的程式碼，其餘命名方式請依此類推。

第一題

根據題目給定的數據，我們可以依據posterior probability由高到低做排序，並且依次計算出某筆數據之下 # of true positive instances (TP)、# of false positive instances (FP)、 # of true negative instances (TN) 、# of false negative instances (FN)、false positive rate (FPR)、true positive rate (TPR)等各項數值(詳細請見表格一)，最後繪製成FPR -TPR平面圖即可得到ROC curve(圖一)。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Rank | Label | TP | FP | TN | FN | FPR | TPR |
| 1 | P | 1 | 0 | 10 | 9 | 0.0 | 0.1 |
| 2 | P | 2 | 0 | 10 | 8 | 0.0 | 0.2 |
| 3 | P | 3 | 0 | 10 | 7 | 0.0 | 0.3 |
| 4 | P | 4 | 0 | 10 | 6 | 0.0 | 0.4 |
| 5 | N | 4 | 1 | 9 | 6 | 0.0 | 0.4 |
| 6 | P | 5 | 1 | 9 | 5 | 0.1 | 0.5 |
| 7 | P | 6 | 1 | 9 | 4 | 0.1 | 0.6 |
| 8 | N | 6 | 2 | 8 | 4 | 0.1 | 0.6 |
| 9 | P | 7 | 2 | 8 | 3 | 0.2 | 0.7 |
| 10 | N | 7 | 3 | 7 | 3 | 0.2 | 0.7 |
| 11 | P | 8 | 3 | 7 | 2 | 0.3 | 0.8 |
| 12 | N | 8 | 4 | 6 | 2 | 0.3 | 0.8 |
| 13 | P | 9 | 4 | 6 | 1 | 0.4 | 0.9 |
| 14 | N | 9 | 5 | 5 | 1 | 0.4 | 0.9 |
| 15 | N | 9 | 6 | 4 | 1 | 0.5 | 0.9 |
| 16 | N | 9 | 7 | 3 | 1 | 0.6 | 0.9 |
| 17 | N | 9 | 8 | 2 | 1 | 0.7 | 0.9 |
| 18 | P | 10 | 8 | 2 | 0 | 0.8 | 1.0 |
| 19 | N | 10 | 9 | 1 | 0 | 0.9 | 1.0 |
| 20 | N  Table 1 Ranking Method to Draw ROC curve and Calculate Area Under Curve | 10 | 10 | 0 | 0 | 1.0 | 1.0 |

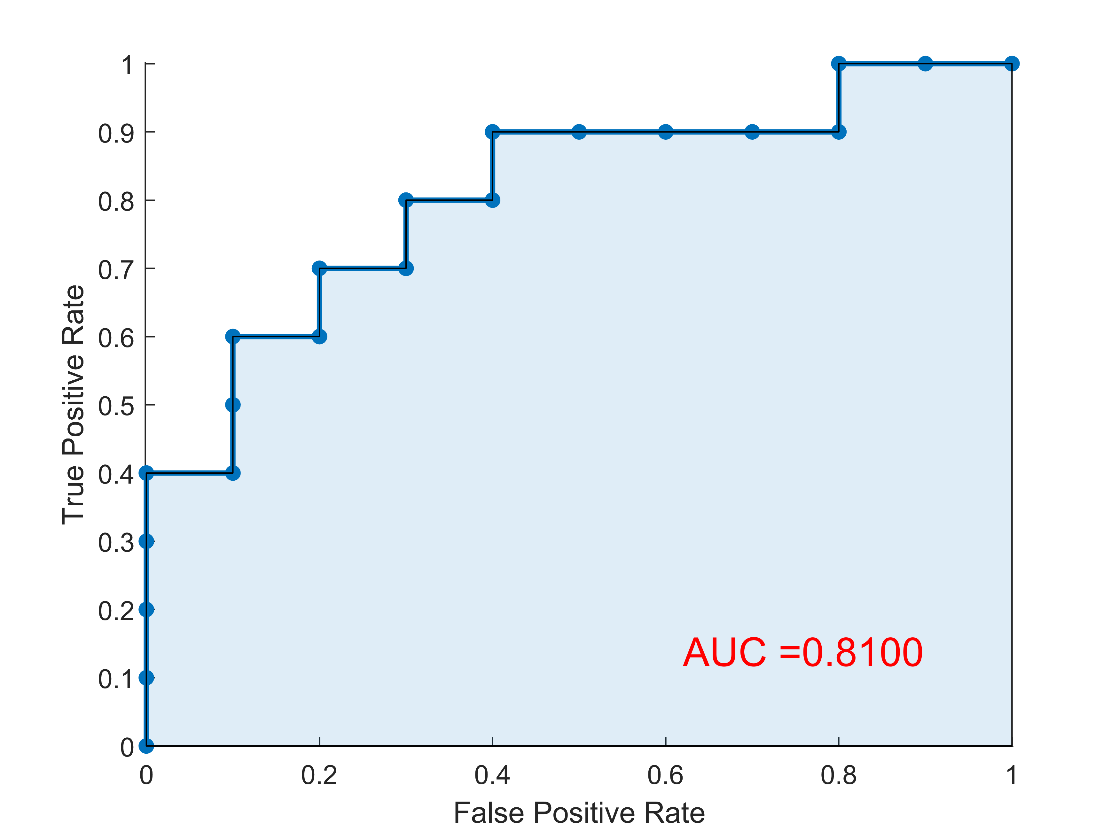


Figure 1 ROC curve to Question 1 and Its Area Under Curve

第二題

利用肉眼計算，我們可以很容易得知： 的最小值發生在 且 。

1. 利用我撰寫的MATLAB code「Q2\_a.m」，總共經過9731次疊代收斂在：

、。

以下表格二列出首三輪與尾三輪的詳細 疊代數值與變化量：

Table 2 Detailed Info. about First Three and Last Three Iterations (Gradient Descent)

|  |  |  |  |
| --- | --- | --- | --- |
|  | |  |  |
|  |  |  | - |
|  |  |  |  |
|  |  |  |  |
| ．．． | | | |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. 利用我撰寫的MATLAB code「Q2\_b.m」，總共經過1次疊代收斂在：

、**。**

以下表格三列出全部的詳細 疊代數值與變化量：

Table 3 Detailed Information about All Iterations (Newton’s Method)

1. sdf;sdlfk;sdkf

根據題目給定的數據，我們可以依據posterior probability由高到低做排序，並且依次計算出某筆數據之下 # of true positive instances (TP)、# of false positive instances (FP)、 # of true negative instances (TN) 、# of false negative instances (FN)、false positive rate (FPR)、true positive rate (TPR)等各項數值(詳細請見表格一)，最後繪製成FPR -TPR平面圖即可得到ROC curve(圖一)。

11

1. 經過3-nearest neighbor algorithm運算，classification結果如下：

|  |  |  |  |
| --- | --- | --- | --- |
| Part | Part | | Correctness |
| Predicted Class | Actual Class |
|  | negative | positive | W |
|  | negative | positive | W |
|  | negative | positive | W |
|  | negative | positive | W |
|  | negative | positive | W |
|  | positive | negative | W |
|  | positive | negative | W |
|  | positive | negative | W |
|  | positive | negative | W |

1. 3-nearest neighbor algorithm的confusion matrix為：

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Predicted Class | |
| ＋ | － |
| Actual Class | ＋ | True  Positive | False  Negative |
| － | False  Positive | True  Negative |

第二題

證明： 且 是對稱正定矩陣 的所有特徵值均為正數

* where is and eigenvalue of

and such that

* 接著，我們考慮 的值：

因為 是real symmetric positive definite matrix，所以

綜合和的結果，我們得到：

* 必為一個正實數

第三題

證明還未完成

第四題

1. 20次的隨機實驗結果分別如下：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Average | No. | Average | No. | Average | No. | Average |
| 1 | 0.5420 | 6 | 0.4857 | 11 | 0.5412 | 16 | 0.5271 |
| 2 | 0.4037 | 7 | 0.4409 | 12 | 0.6227 | 17 | 0.3614 |
| 3 | 0.4785 | 8 | 0.6096 | 13 | 0.6752 | 18 | 0.5278 |
| 4 | 0.4315 | 9 | 0.4478 | 14 | 0.3314 | 19 | 0.5932 |
| 5 | 0.3910 | 10 | 0.4941 | 15 | 0.5567 | 20 | 0.5819 |

1. 50次的隨機實驗結果分別如下：

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Average | No. | Average | No. | Average | No. | Average | No. | Average |
| 1 | 0.4946 | 11 | 0.5166 | 21 | 0.4986 | 31 | 0.5062 | 41 | 0.4830 |
| 2 | 0.4925 | 12 | 0.5030 | 22 | 0.5123 | 32 | 0.5010 | 42 | 0.5120 |
| 3 | 0.5035 | 13 | 0.5002 | 23 | 0.5230 | 33 | 0.4925 | 43 | 0.4958 |
| 4 | 0.4998 | 14 | 0.5017 | 24 | 0.4798 | 34 | 0.4907 | 44 | 0.4963 |
| 5 | 0.5031 | 15 | 0.4938 | 25 | 0.5163 | 35 | 0.4997 | 45 | 0.5076 |
| 6 | 0.5153 | 16 | 0.5139 | 26 | 0.5006 | 36 | 0.5086 | 46 | 0.4936 |
| 7 | 0.4895 | 17 | 0.5106 | 27 | 0.5116 | 37 | 0.5114 | 47 | 0.4982 |
| 8 | 0.4936 | 18 | 0.5079 | 28 | 0.4950 | 38 | 0.4889 | 48 | 0.5017 |
| 9 | 0.4879 | 19 | 0.5099 | 29 | 0.4999 | 39 | 0.5065 | 49 | 0.4968 |
| 10 | 0.5032 | 20 | 0.4991 | 30 | 0.5061 | 40 | 0.4942 | 50 | 0.4888 |

第五題

1. 取 、、，則原方程組即為
2. 求解normal equation：

、

利用Gauss-Jordan消去法，我們得到以下結果：

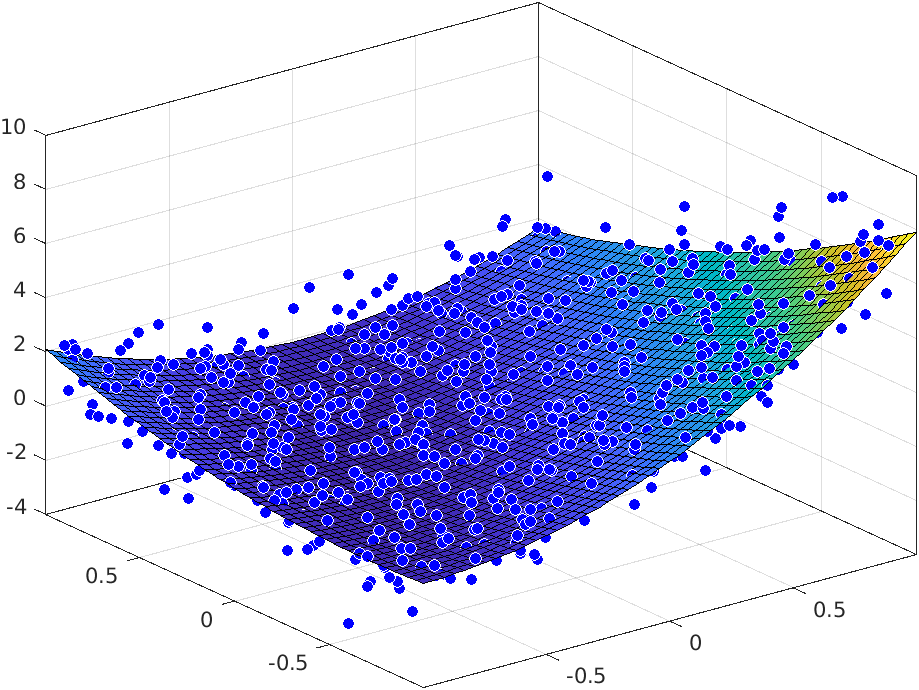
1. 因此，least squares approximation solution

第六題

1. 關於本題使用到的training dataset ，詳細請見[GitHub](https://github.com/a2468834/Machine_Learning_Grad/blob/master/HW1/MATLAB%20code/Q6_dataset.csv)連結。
2. 利用MATLAB內建函數fit()，我們可以得到以下的fitting surface：

1. 先用 計算出input data的預測值 ，其與實際值 的差異即為error；接著將1000筆資料的error取絕對值再平均，即可求得MAE。

，每筆資料的error值詳細請見[GitHub](https://github.com/a2468834/Machine_Learning_Grad/blob/master/HW1/MATLAB%20code/Q6_error_vector.csv)連結。

另外，這個fitting surface的3D圖如下：