UMC 205 Assignment I

Mehul Shrivastava

January 2024

Q1:

1.a

1.a (1 mark) The parameters for your model are $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1$. Using assumption (A1), what can you say about $\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1$?.

On assuming that each of the vector is mutually independent, we can say that $\sum_0$ & $\sum_1$ are diagonal matrices.

1.b

1.b (1 mark) State the Bayes classifier under the modified loss function.

Given: l(0,1) = 5 and l(1,0) = 1
Assume: $p_0 = P(Y = 0)$ and $p_1 = P(Y = 1)$ We can assume them 1/2 as number of sample are equal
Assume: $n_k(x) = P(\frac{Y=k}{X=x})$ k = 0 or 1

The condition for this looks like:
The Bayes Classifier for this under the modified loss function is:
we'll select Y=1 if $5 * n_1(x) > n_0(x)$ using the modified loss function.

$= log(5 * n_1(x)) > log(n_0(x))$

$= log(5) - \frac{1}{2} * (x - u_1)^T \sum_1 (x - u_1) - \frac{1}{2} log(|\sum_1|) > \frac{-1}{2}(x - u_0)^T \sum_0 (x - u_0) - \frac{1}{2} log(|\sum_0|)$
(assuming p_0 and p_1 are equal to 1/2)

$= log(5) - \frac{1}{2} * (x - u_1)^T \sum_1 (x - u_1) - \frac{1}{2} log(|\sum_1|) + \frac{1}{2}(x - u_0)^T \sum_0 (x - u_0) + \frac{1}{2} log(|\sum_0|) > 0$

Let the above inequality be f(x), then the bayes classifier $h^*(x)$can be written as:

$$\begin{cases} 1 & f(x) > 0 \\ 0 & f(x) < 0 \end{cases}$$

1.c

1.c (2 marks) Obtain sample estimates of the class conditional means and variances for your model with $n = 2, 10, 20, 50, 100, 500, 1000$ samples. Write down the estimates for all the parameters in a table for different values of n.

As, the question is only asking about variances. I made the non diagonal entries in the Co-variance matrix to be 0

1 n=2

1.1 Class 0 Mean and Covariance Matrix Class 0:

Mean : $[-1.14479, -1.99227, -0.877605, -1.083265, -2.46712]$

$$\text{Cov:} \begin{bmatrix} 2.96912180e-03 & 0 & 0 & 0 & 0 \\ 0 & 8.40614804e+00 & 0 & 0 & 0 \\ 0 & 0 & 1.37267537e+00 & 0 & 0 \\ 0 & 0 & 0 & 1.56884011e-01 & 0 \\ 0 & 0 & 0 & 0 & 5.90610593e-01 \end{bmatrix}$$

So, the variance is $[2.96912180e-03, 8.40614804e+00, 1.37267537e+00, 1.56884011e-01, 5.90610593e-01]$

1.2 Class 1 :Mean and Covariance Matrix

$$\text{Mean : } [-0.151365, 0.877535, -0.429675, 0.964285, 1.937245]$$

$$\text{Covariance Matrix : } \begin{bmatrix} 4.95311045e-03 & 0 & 0 & 0 & 0 \\ 0 & 2.24826012e-02 & 0 & 0 & 0 \\ 0 & 0 & 1.38228564e-02 & 0 & 0 \\ 0 & 0 & 0 & 3.76018691e+00 & 0 \\ 0 & 0 & 0 & 0 & 7.81480438e+00 \end{bmatrix}$$

$$\text{So, the variance is } [4.95311045e-03, 2.24826012e-02, 1.38228564e-02, 3.76018691e+00, 7.81480438e+00]$$

2 n=10

2.1 Class 0:

Mean: [-1.28809917, -1.04324333, -0.4994225, -0.98741583, -0.34477583]

$$\text{Covariance Matrix: } \begin{bmatrix} 0.96152185 & 0 & 0 & 0 & 0 \\ 0 & 3.02206183 & 0 & 0 & 0 \\ 0 & 0 & 4.1461748 & 0 & 0 \\ 0 & 0 & 0 & 1.79324502 & 0 \\ 0 & 0 & 0 & 0 & 5.34836002 \end{bmatrix}$$

Class 0 variance Matrix: [0.96152185, 3.02206183, 4.1461748, 1.79324502, 5.34836002]

2.2 Class 1:

$$\text{Mean: } [1.5466625, 0.57448083, 0.5426725, 1.247495, 1.50729667]$$

$$\text{Covariance Matrix: } \begin{bmatrix} 2.76932331 & 0 & 0 & 0 & 0 \\ 0 & 1.5648745 & 0 & 0 & 0 \\ 0 & 0 & 3.09023413 & 0 & 0 \\ 0 & 0 & 0 & 1.0234835 & 0 \\ 0 & 0 & 0 & 0 & 6.79049305 \end{bmatrix}$$

Class 1 Variance Matrix: [2.76932331, 1.5648745, 3.09023413, 1.0234835, 6.79049305]

3 n=20

3.1 Class 0:

Mean: [-1.22665562, -1.13424375, -0.95430938, -0.93542594, -0.70939094]

$$\text{Covariance Matrix: } \begin{bmatrix} 1.22545418 & 0 & 0 & 0 & 0 \\ 0 & 2.52422032 & 0 & 0 & 0 \\ 0 & 0 & 2.59160887 & 0 & 0 \\ 0 & 0 & 0 & 2.23580914 & 0 \\ 0 & 0 & 0 & 0 & 3.79642161 \end{bmatrix}$$

Class 0 Variance Matrix: [1.22545418, 2.52422032, 2.59160887, 2.23580914, 3.79642161]

3.2 Class 1:

$$\text{Mean: } [1.38296938, 0.73431531, 0.92642594, 1.19777656, 1.65929875]$$

$$\text{Covariance Matrix: } \begin{bmatrix} 2.33147076 & 0 & 0 & 0 & 0 \\ 0 & 1.91455818 & 0 & 0 & 0 \\ 0 & 0 & 3.74006288 & 0 & 0 \\ 0 & 0 & 0 & 1.24287538 & 0 \\ 0 & 0 & 0 & 0 & 6.16691229 \end{bmatrix}$$

Class 1 Variance Matrix: [2.33147076, 1.91455818, 3.74006288, 1.24287538, 6.16691229]

4 n=50

4.1 Class 0:

Mean: [-1.04311439, -1.00680598, -0.83114293, -0.87435768, -0.92853329]

$$\text{Covariance Matrix:} \begin{bmatrix} 2.46054533 & 0 & 0 & 0 & 0 \\ 0 & 2.1551572 & 0 & 0 & 0 \\ 0 & 0 & 2.27656248 & 0 & 0 \\ 0 & 0 & 0 & 2.50437093 & 0 \\ 0 & 0 & 0 & 0 & 4.17718552 \end{bmatrix}$$

Class 0 Variance Matrix: [2.46054533, 2.1551572, 2.27656248, 2.50437093, 4.17718552]

4.2 Class 1:

$$\text{Mean: } [1.08938707, 0.77834756, 0.94893768, 1.17204854, 1.29407963]$$

$$\text{Covariance Matrix:} \begin{bmatrix} 2.22608175 & 0 & 0 & 0 & 0 \\ 0 & 2.2317375 & 0 & 0 & 0 \\ 0 & 0 & 3.14358394 & 0 & 0 \\ 0 & 0 & 0 & 1.96452843 & 0 \\ 0 & 0 & 0 & 0 & 5.65558063 \end{bmatrix}$$

Class 1 Variance Matrix: [2.22608175, 2.2317375, 3.14358394, 1.96452843, 5.65558063]

5 n=100

5.1 Class 0:

Mean: [-0.99476846, -0.96928066, -1.05481621, -0.94446868, -1.08280835]

$$\text{Covariance Matrix:} \begin{bmatrix} 2.19778867 & 0 & 0 & 0 & 0 \\ 0 & 2.20489789 & 0 & 0 & 0 \\ 0 & 0 & 2.25184337 & 0 & 0 \\ 0 & 0 & 0 & 3.36463841 & 0 \\ 0 & 0 & 0 & 0 & 4.5246874 \end{bmatrix}$$

Class 0 Variance Matrix: [2.19778867, 2.20489789, 2.25184337, 3.36463841, 4.5246874]

5.2 Class 1:

$$\text{Mean: } [1.03363769, 0.86751989, 1.03496885, 1.16551324, 1.13017736]$$

$$\text{Covariance Matrix:} \begin{bmatrix} 2.45626728 & 0 & 0 & 0 & 0 \\ 0 & 2.21291918 & 0 & 0 & 0 \\ 0 & 0 & 2.7606818 & 0 & 0 \\ 0 & 0 & 0 & 2.28202453 & 0 \\ 0 & 0 & 0 & 0 & 5.19667639 \end{bmatrix}$$

Class 1 Variance Matrix: [2.45626728, 2.21291918, 2.7606818, 2.28202453, 5.19667639]

6 n=500

6.1 Class 0:

Mean: [-0.99137334, -1.01002867, -1.0135317, -0.95962183, -0.98810268]

$$\text{Covariance Matrix: } \begin{bmatrix} 2.41358857 & 0 & 0 & 0 & 0 \\ 0 & 2.17629641 & 0 & 0 & 0 \\ 0 & 0 & 2.44973644 & 0 & 0 \\ 0 & 0 & 0 & 3.95741295 & 0 \\ 0 & 0 & 0 & 0 & 5.2336306 \end{bmatrix}$$

Class 0 Variance Matrix: [2.41358857, 2.17629641, 2.44973644, 3.95741295, 5.2336306]

6.2 Class 1:

$$\text{Mean: } [0.94343405, 0.9603001, 0.97646207, 1.01706796, 0.9647982]$$

$$\text{Covariance Matrix: } \begin{bmatrix} 2.42456845 & 0 & 0 & 0 & 0 \\ 0 & 2.42134545 & 0 & 0 & 0 \\ 0 & 0 & 2.42540283 & 0 & 0 \\ 0 & 0 & 0 & 3.29783405 & 0 \\ 0 & 0 & 0 & 0 & 5.23720202 \end{bmatrix}$$

Class 1 Variance Matrix: [2.42456845, 2.42134545, 2.42540283, 3.29783405, 5.23720202]

7 n=1000

7.1 Class 0:

Mean: [-0.96692373, -0.94917209, -1.01474212, -0.9756399, -1.02319919]

$$\text{Covariance Matrix: } \begin{bmatrix} 2.42104548 & 0 & 0 & 0 & 0 \\ 0 & 2.26265116 & 0 & 0 & 0 \\ 0 & 0 & 2.50647719 & 0 & 0 \\ 0 & 0 & 0 & 3.92028292 & 0 \\ 0 & 0 & 0 & 0 & 5.30605245 \end{bmatrix}$$

Class 0 Variance Matrix: [2.42104548, 2.26265116, 2.50647719, 3.92028292, 5.30605245]

7.2 Class 1:

$$\text{Mean: } [0.95342046, 0.97475045, 0.97757304, 1.00629518, 0.97115193]$$

$$\text{Covariance Matrix: } \begin{bmatrix} 2.44145557 & 0 & 0 & 0 & 0 \\ 0 & 2.36649939 & 0 & 0 & 0 \\ 0 & 0 & 2.41741063 & 0 & 0 \\ 0 & 0 & 0 & 3.66478154 & 0 \\ 0 & 0 & 0 & 0 & 5.31449108 \end{bmatrix}$$

Class 1 Variance Matrix: [2.44145557, 2.36649939, 2.41741063, 3.66478154, 5.31449108]

1-d

The misclassification loss for n = 2 is 500. for n = 10 is 66 for n = 20 is 39 for n = 50 is 35 for n = 100 is 37 for n = 500 is 37 for n = 1000 is 46 The Normalized loss can be found by dividing each of them by 200.

n	Misclassification Loss	Normalized Loss
2	500	2.5
10	66	0.33
20	39	0.195
50	35	0.175
100	37	0.185
500	37	0.185
1000	46	0.23

Table 1: Misclassification Loss for Different n

1-e

From running the code, the accuracy is 0.5.

Q2

The first four parts are answered below:

- Weight vector ($\mathbf{w}$): $[32.89384, 32.7266, 32.51577, 24.58485, 15.88242]$
- Number of errors: 1926
- Margin: 0.0006139827770785612
- Radius of the dataset: 4.429983960817466

The next four parts which is based on the MNIST data is:

- Number of errors: 66
- Margin: 13.262076260785593
- Radius of the dataset: 4380.656685931916

Q3:

3.1.a

Normalized projection vector **w**: $[0.28543316, 0.2165812, -0.65799713, -0.66231433]$

Threshold b: -0.7120525349911094

3.1.b

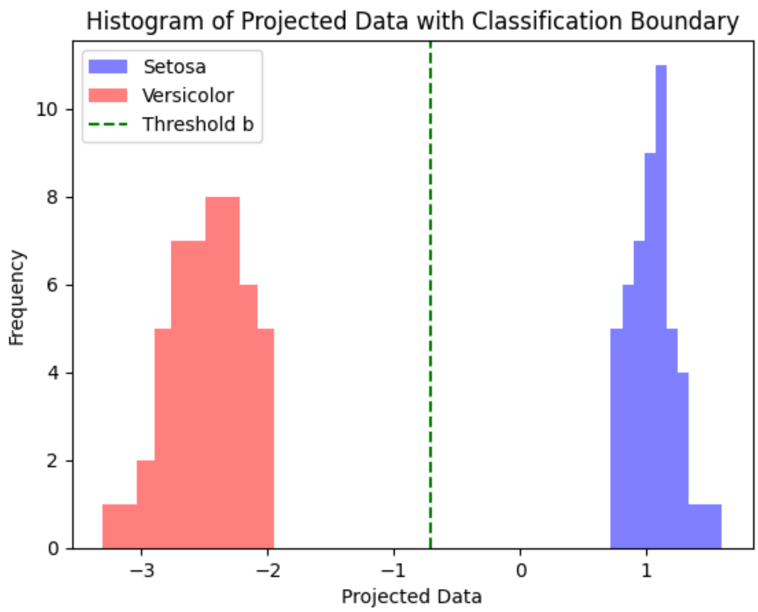

Figure 1: Histogram of the projected data with classifier boundary

7.3 3.2.a

w normalized: [0.745084140.66697048]

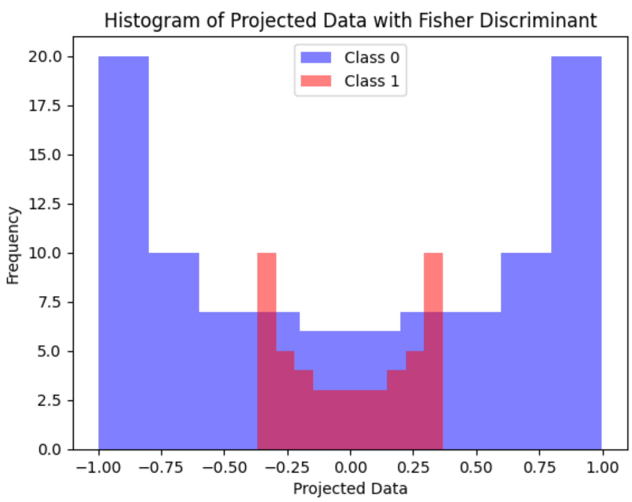

Figure 2: Histogram of the projected data

The data is clearly not linearly separable.

3.2.b

The data is linearly separable.

Threshold b : 0.40142597354547827
Normalized projection vector $w_{\text{normalized}}$: $[-7.98454152 \times 10^{-18}, 8.31770944 \times 10^{-18}, 0.707106781, 0.707106781]$

3.2.c

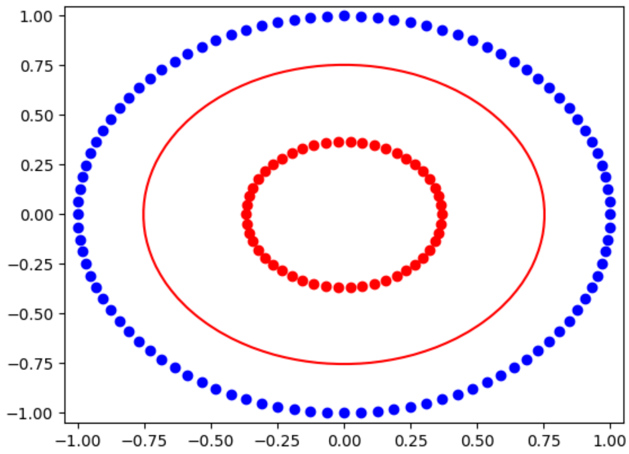

Figure 3: Plot along with classifier boundary

3.2.d

Threshold	Test Accuracy
0.101426	0.67
0.201426	0.75
0.301426	0.80
0.401426	0.77
0.501426	0.75
0.601426	0.75
0.701426	0.75

Table 2: Test Accuracy for Various Thresholds

Q4:

After running the K-fold, the following is the final plot of various metrics:

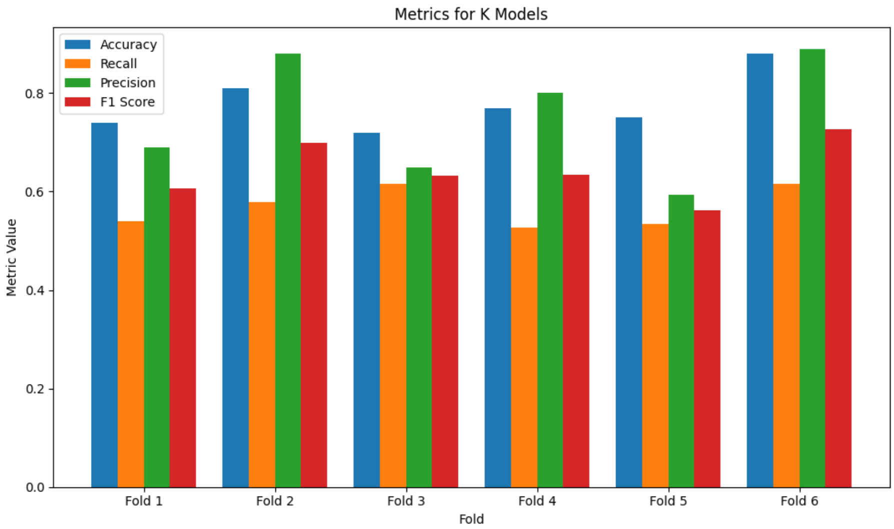

Figure 4: Various metrics

The following is the data in tabular form:

Fold	Accuracy	Recall	Precision	F1 Score
Fold 1	0.74	0.5405	0.6897	0.6061
Fold 2	0.81	0.5789	0.88	0.6984
Fold 3	0.72	0.6154	0.6486	0.6316
Fold 4	0.77	0.5263	0.8	0.6349
Fold 5	0.75	0.5333	0.5926	0.5614
Fold 6	0.88	0.6154	0.8889	0.7273
Average	**0.7783**	**0.5683**	**0.7500**	**0.6433**

Table 3: Metrics Across Folds