1.5 Midterm Redo

Fix problems 2a) & 2b)

2a) choose mean instead of previous answer (mode)

$\hat{B} = \int_0^\infty b e^{-b}\, db$ let $u = x$ $du = dx$

$dv = e^{-x} dx$ $v = -e^{-x}$

$\int u\,dv = uv - \int v\,du$

$= -xe^{-x} + \int e^{-x} dx = -xe^{-x} + e^{-x} \Big|_0^\infty = \boxed{1}$

2b) $\hat{B}_{map}$

maximized $f(y_{min} | b)\, f_B(b)$

$f(y_{min} | b)\, f_B(b) = \begin{cases} 1 - v^0 (e^{-y_{min}}), & b = y_{min} \\ v^0 (e^{-y_{min}+1}), & b = y_{min} - 1 \end{cases}$

$(1-v^0) e^{-y_{min}} \underset{y_{min}-1}{\overset{y_{min}}{\gtrless}} v^0 e^{-y_{min}+1}$

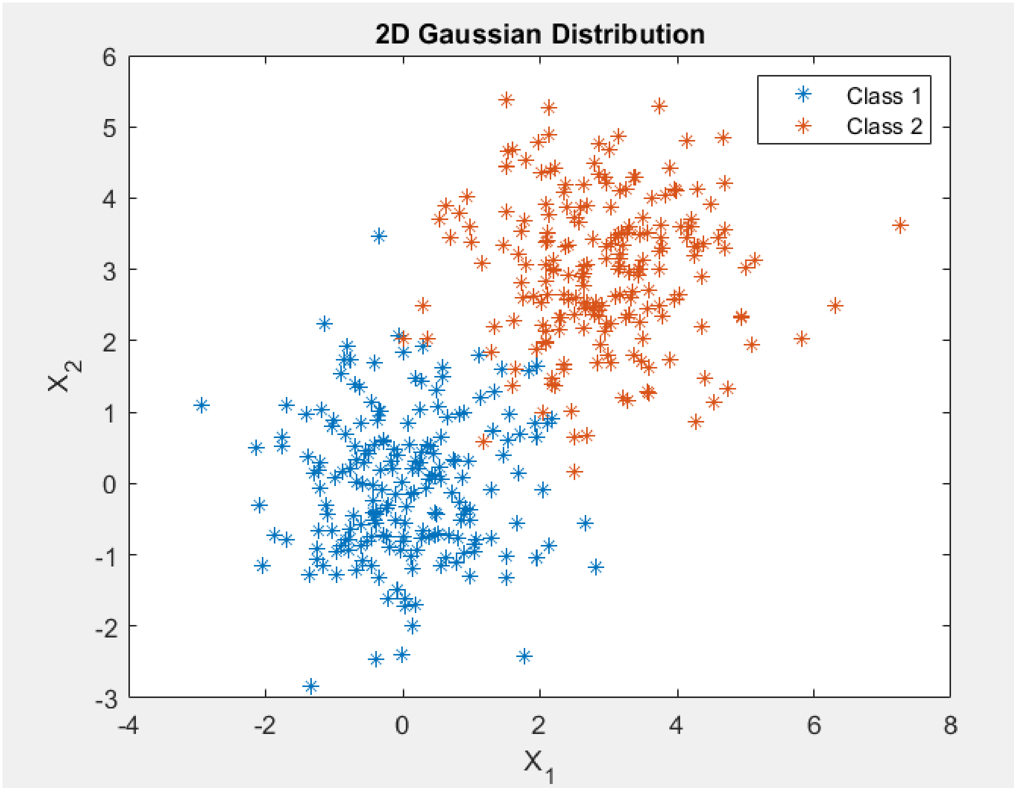
$\cancel{e^{-y_{min}}} - v^0 \cancel{e^{-y_{min}}} \gtrless v^0 (e \cdot \cancel{e^{-y_{min}}})$

$1 - v^0 \gtrless v^0 e$

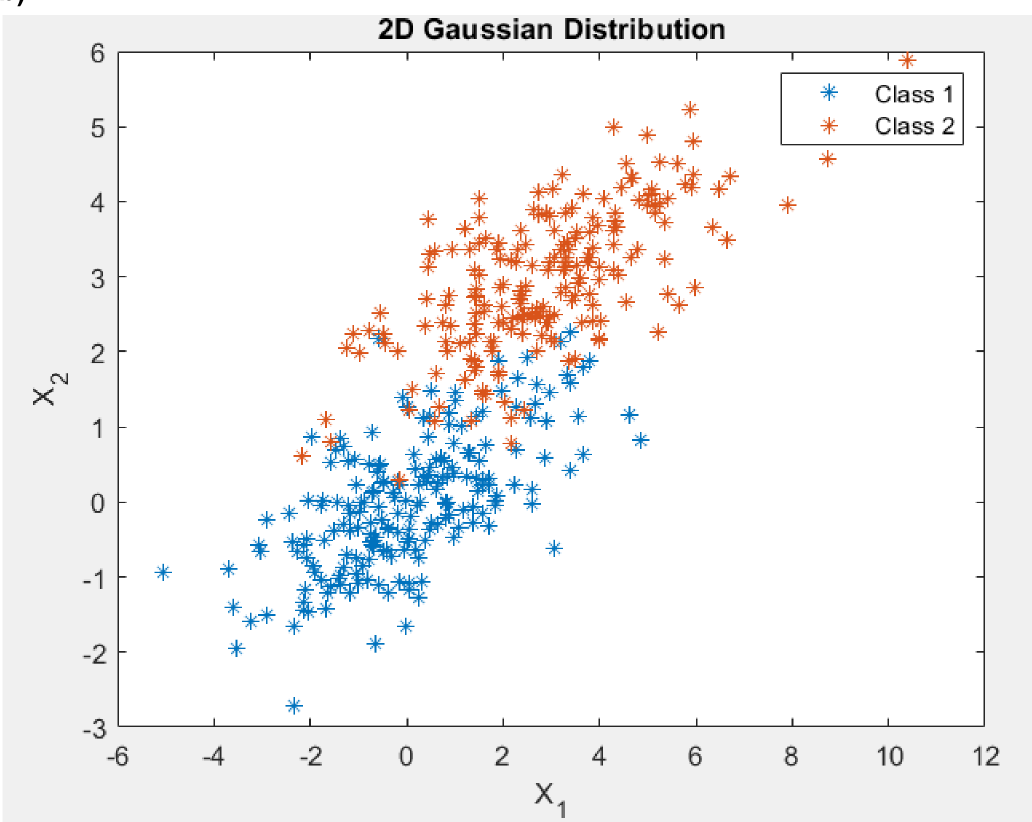
$1 \gtrless v^0 (e+1)$

$\boxed{\frac{1}{e+1} \underset{\text{choose } y_{min}-1}{\overset{\text{choose } y_{min}}{\gtrless}} v^0}$

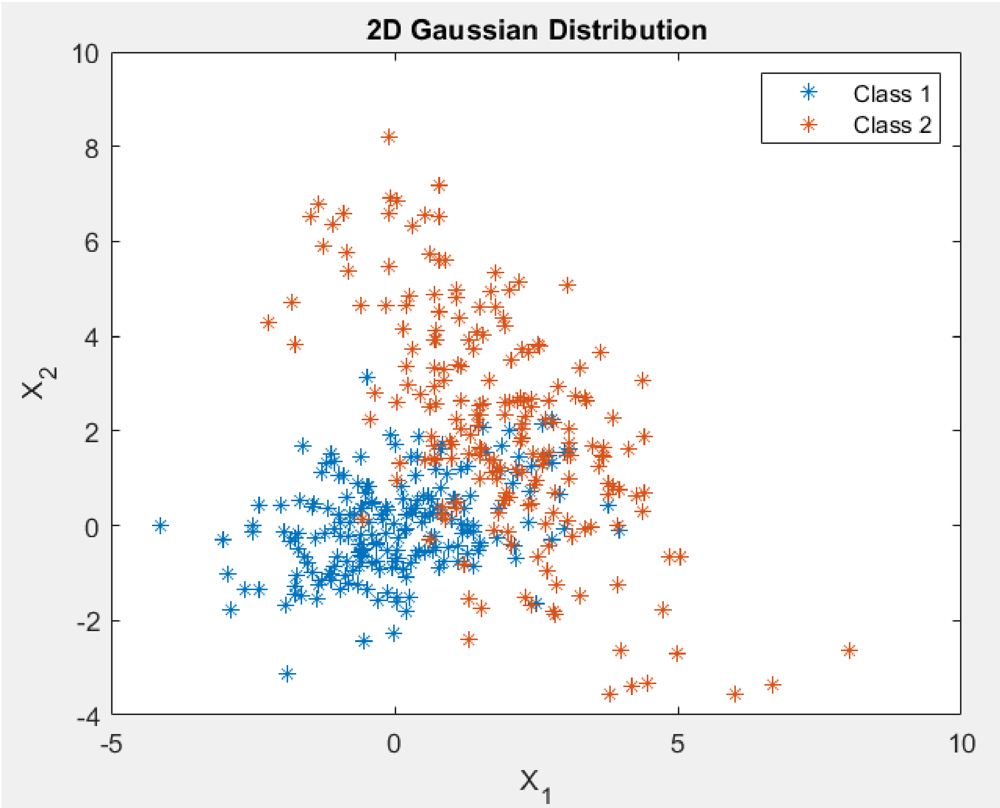
1.6a)


2D Gaussian Distribution
Class 1
Class 2
6
5
4
3
2
1
0
-1
-2
-3
-4
-2
0
2
4
6
8
X_2
X_1

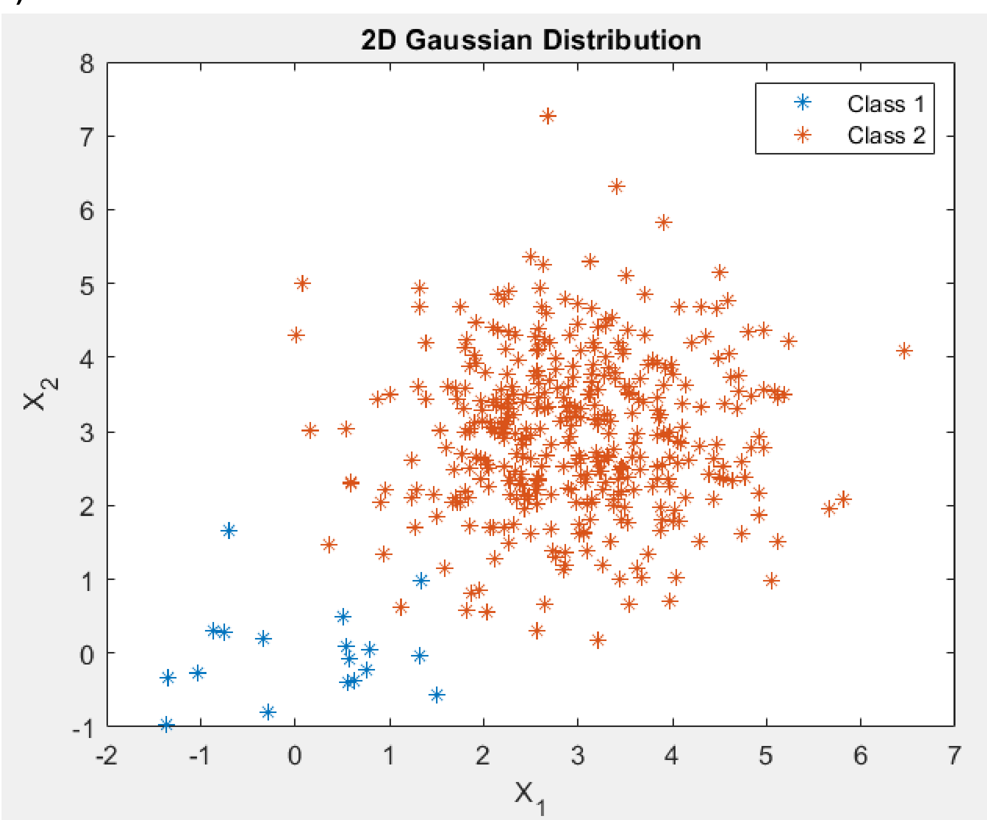
b)


2D Gaussian Distribution
Class 1
Class 2
6
5
4
3
2
1
0
-1
-2
-3
-6
-4
-2
0
2
4
6
8
10
12
X_2
X_1

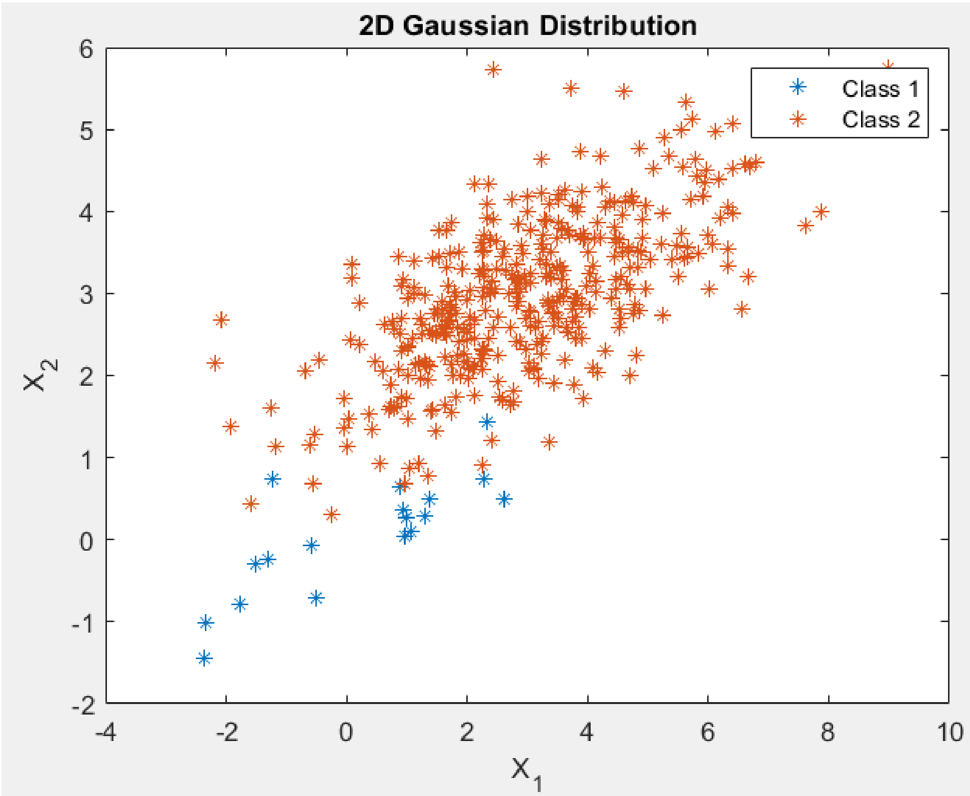
c)


2D Gaussian Distribution
Class 1
Class 2
X2
X1
10
8
6
4
2
0
-2
-4
-5
0
5
10

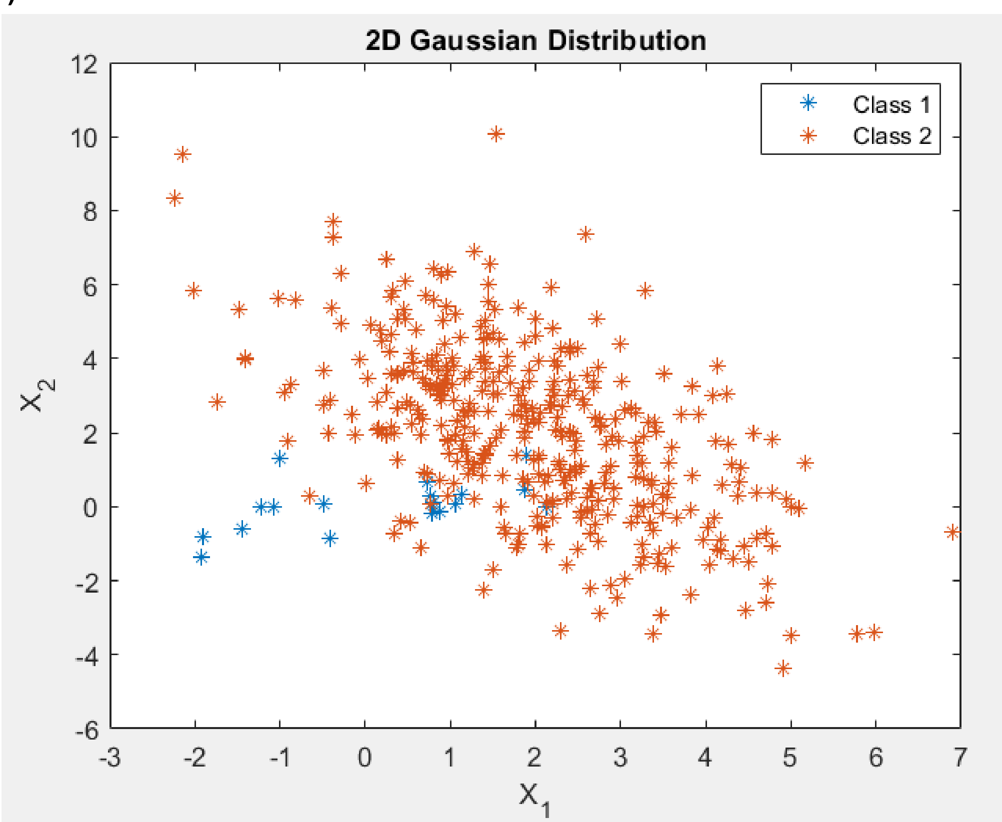
d)


2D Gaussian Distribution
Class 1
Class 2
X2
X1
8
7
6
5
4
3
2
1
0
-1
-2
-1
0
1
2
3
4
5
6
7

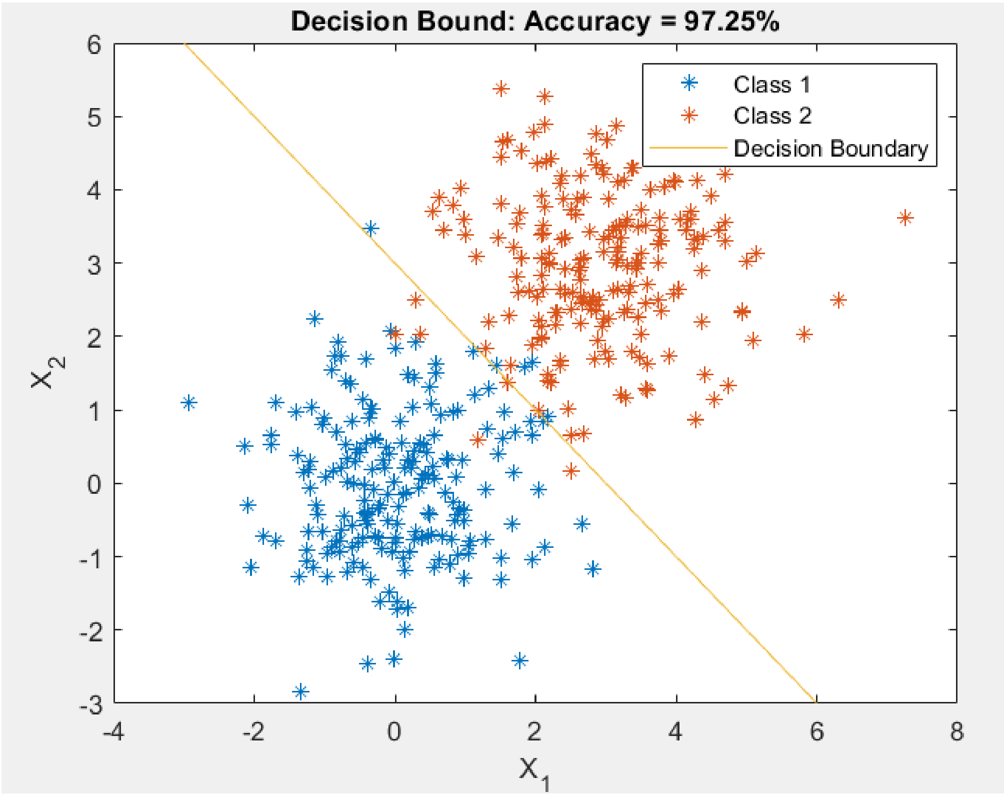
e)


2D Gaussian Distribution
Class 1
Class 2
X_2
X_1

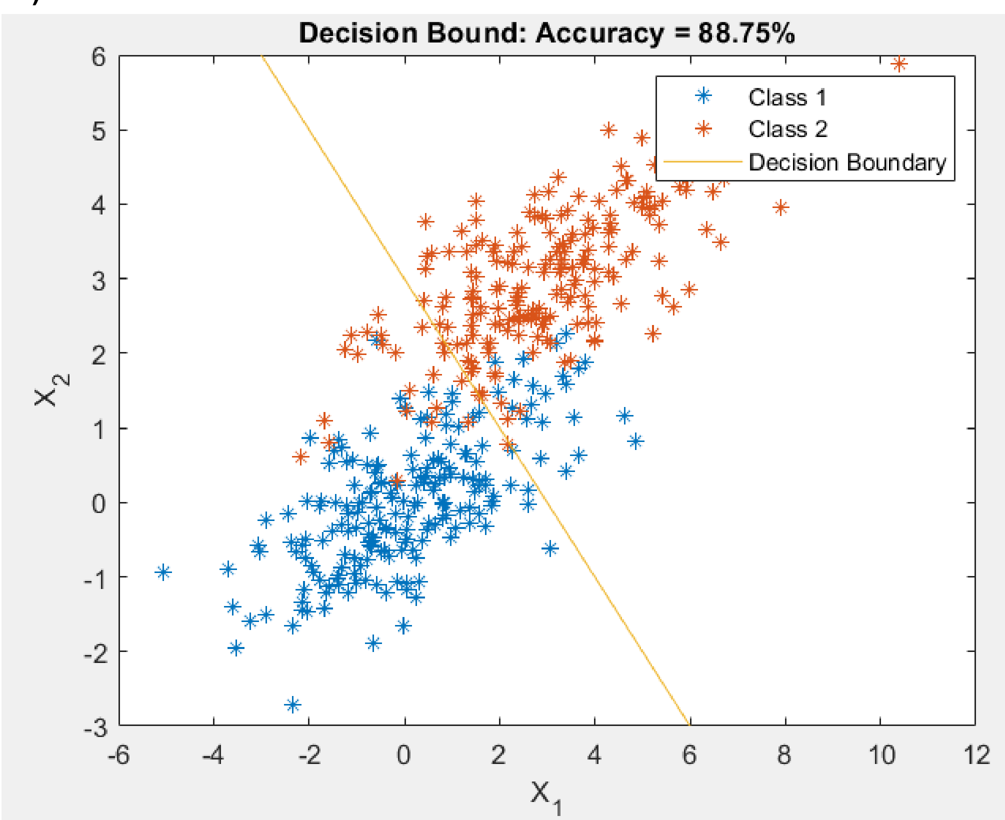
f)


2D Gaussian Distribution
Class 1
Class 2
X_2
X_1

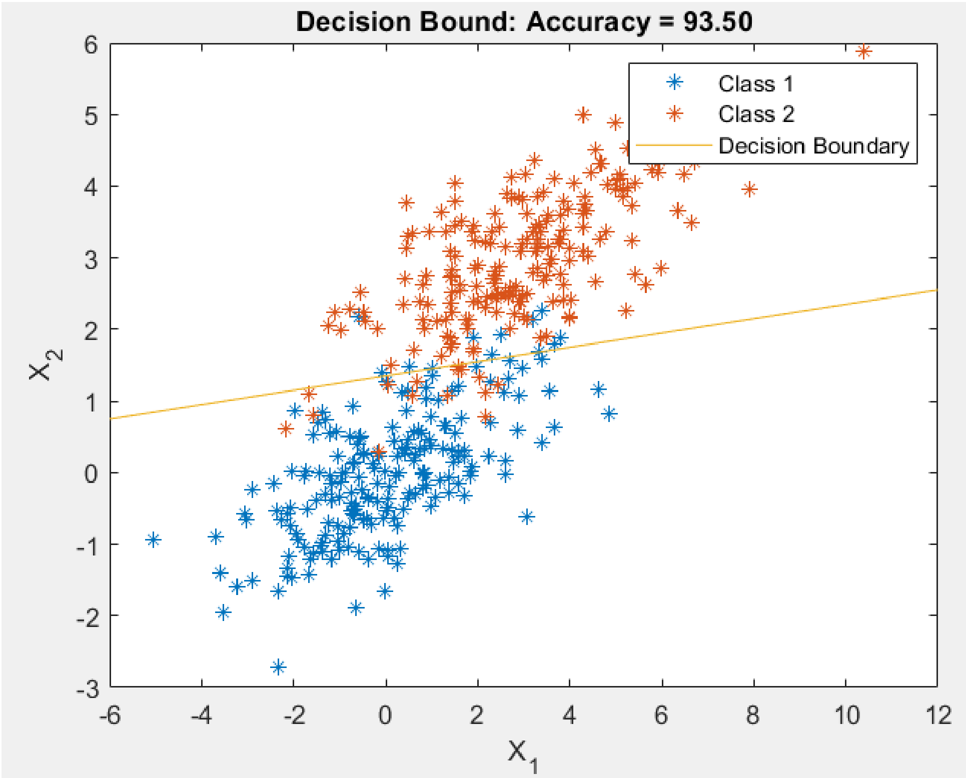
1.7a) Case 1



b1) Case 1

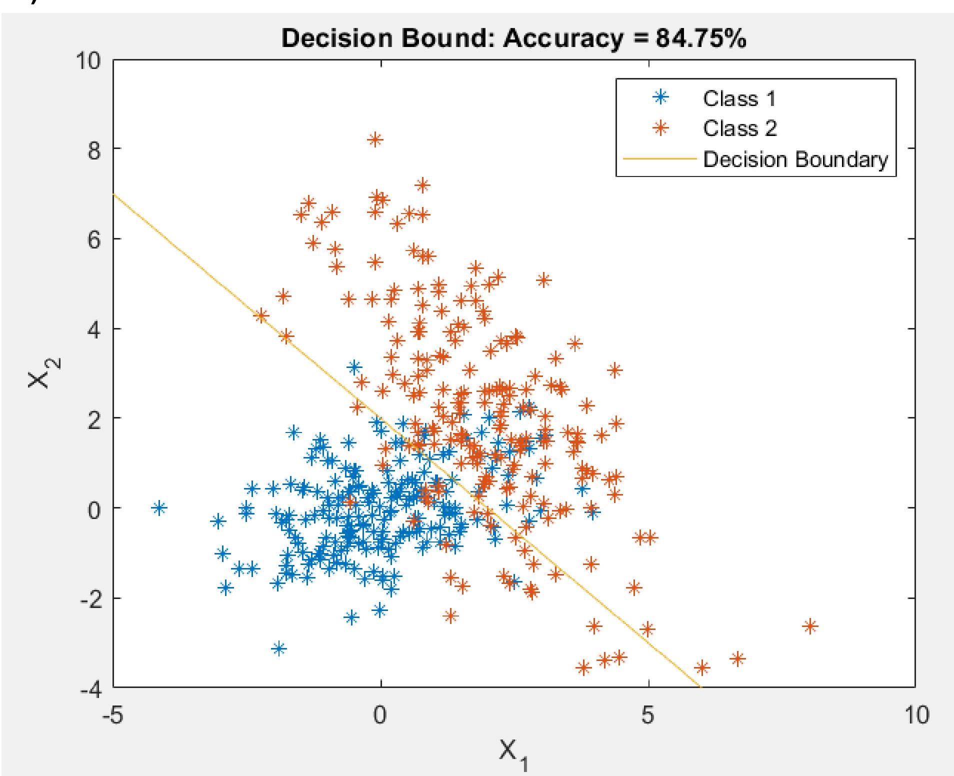


b2) Case 2

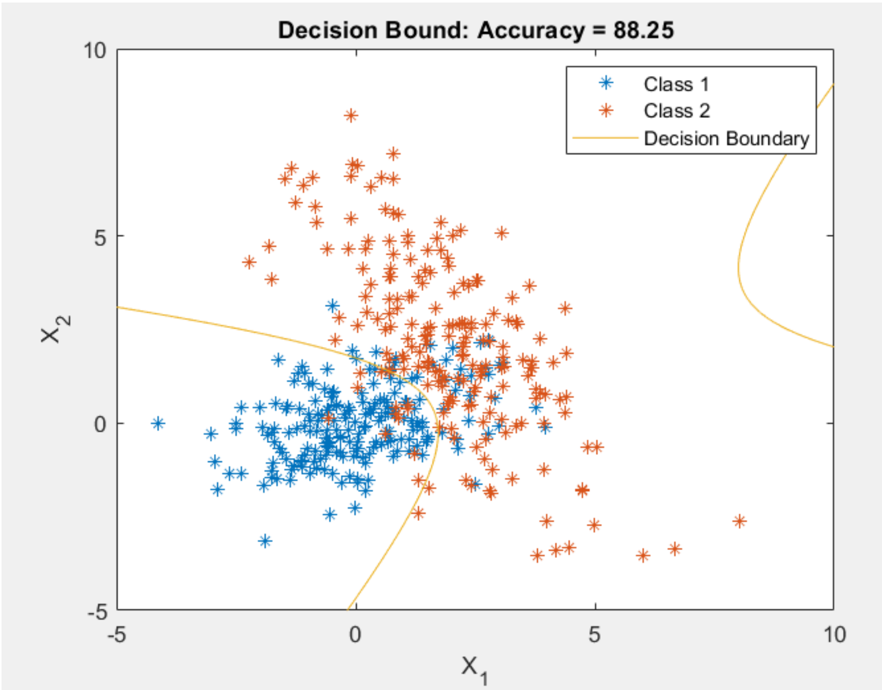


b) There IS a difference between the two above plots. Case 1 does not consider a non-identity distribution while case 2 does. This makes case 2 more accurate.

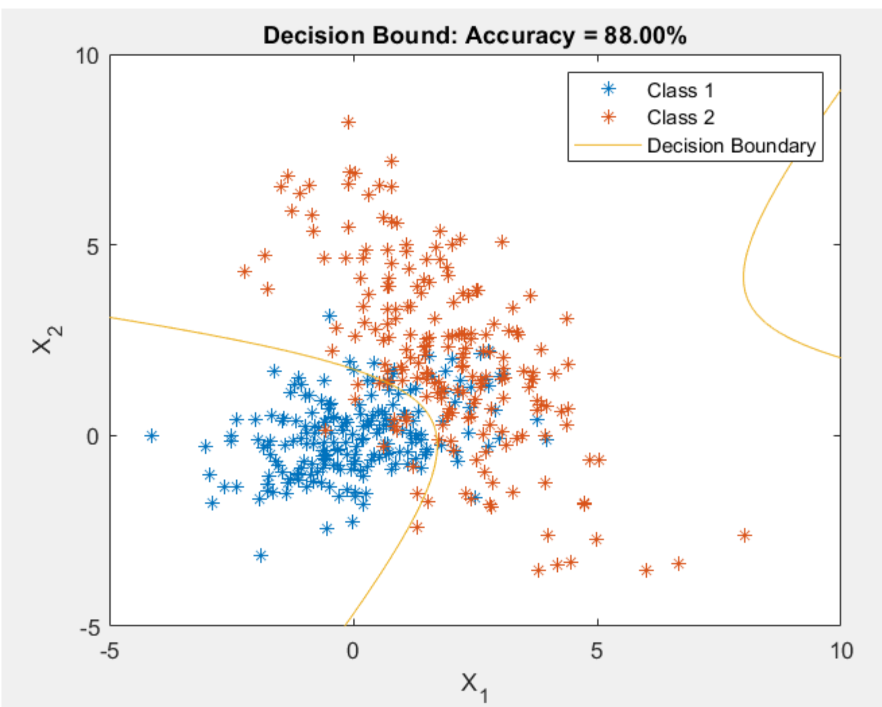
c1) Case 1



c2) Case 2

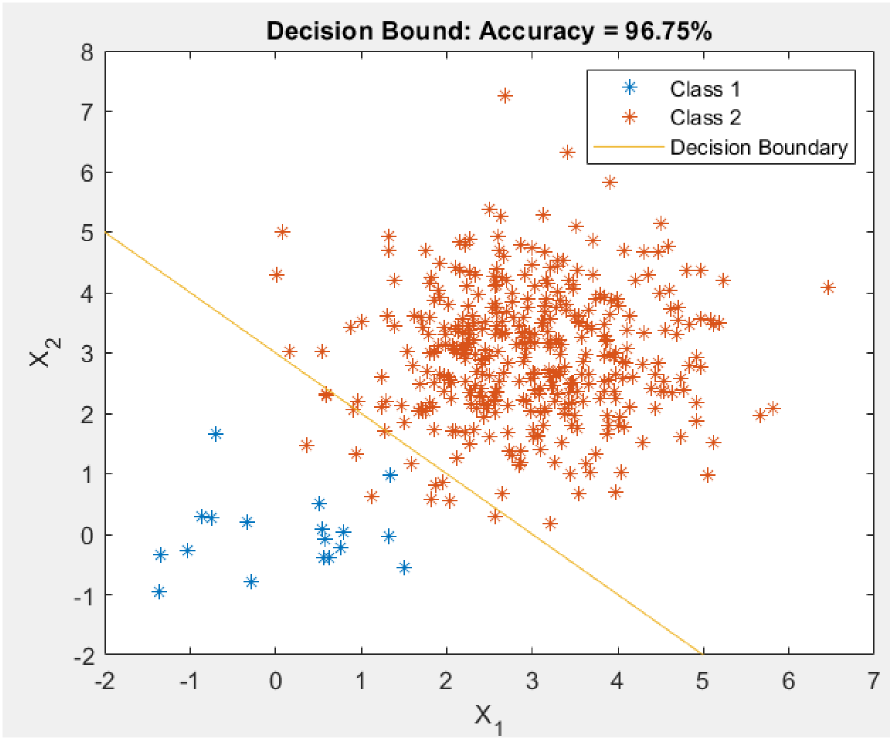


c3) Case 3



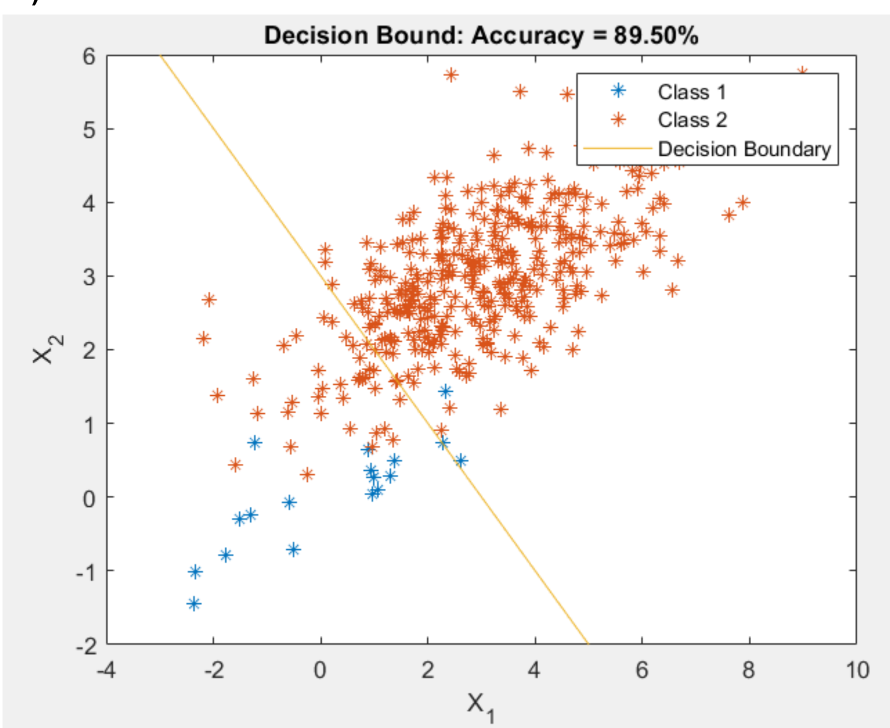
c) There is a significant difference in the plots for part c. The first is linear, while the other two are nonlinear. The first is also significantly less accurate than the other two decision boundaries. Case 2 and Case 3 are very similar in accuracy, so I'd recommend generating and testing more data before deciding which is more effective. Intuitively, Case 3 should have a better prediction as it considers more info than Case 2 (the differing values of sigma), but more data is needed to be sure.

d) Case 1

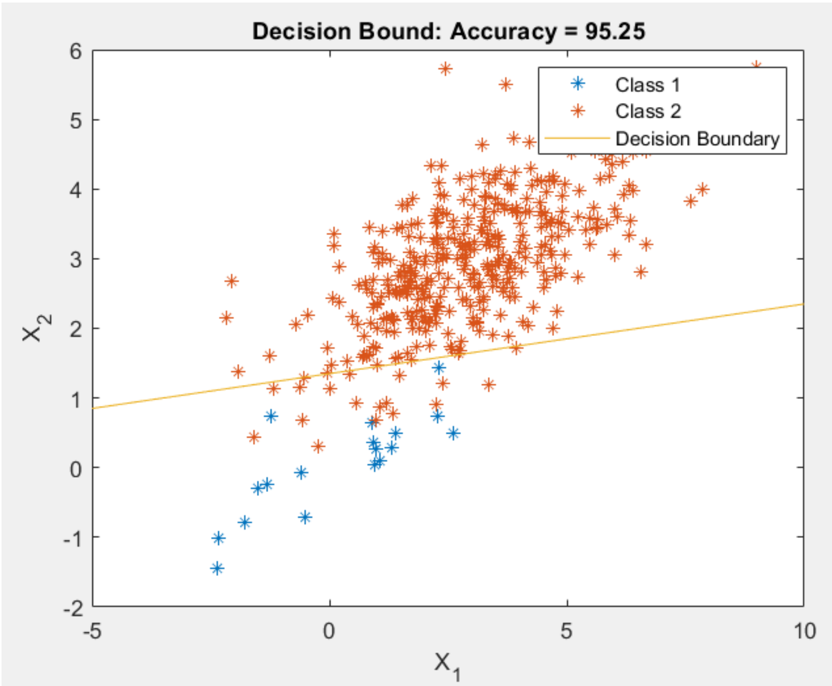


d) If we just used the higher prior to make our decision the model would be 95% accurate. This is less than the 96.75% of our real decision boundary. But regardless of the percentage change, this model would be very bad since we would never guess the rarer class. Accuracy is sometimes deceiving.

e1) Case 1

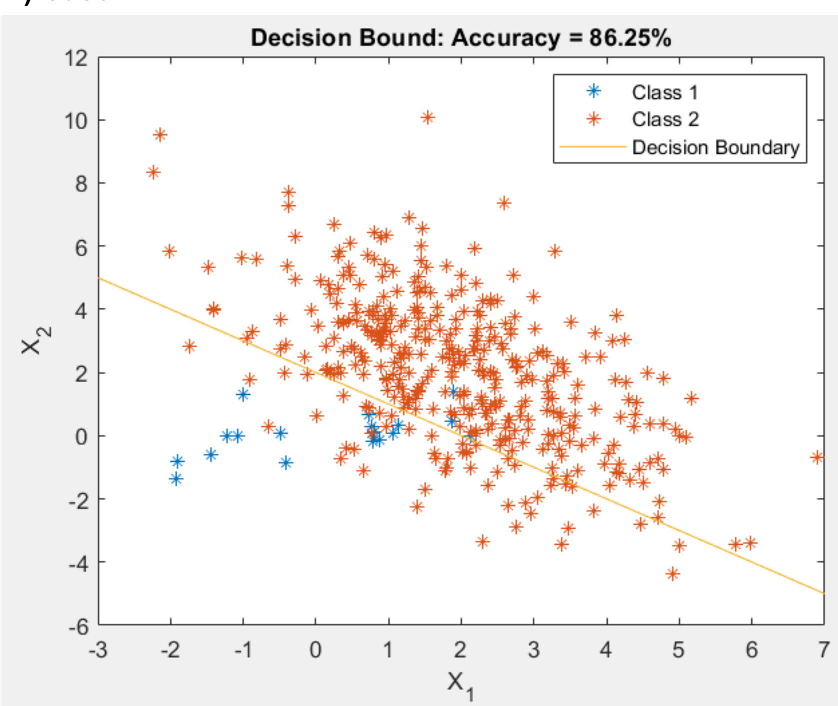


e2) Case 2

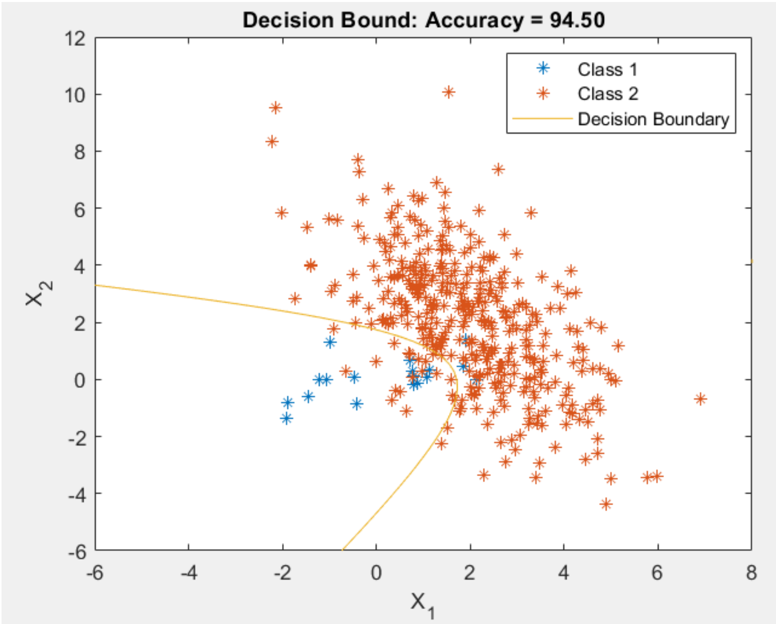


e) There IS a difference between the two above plots. Case 1 does not consider a non-identity distribution while case 2 does. This makes case 2 (much) more accurate. If we just used the higher prior to make our decision the model would be 95% accurate. This is slightly less than the 95.25% of our better decision boundary and more than our worst (%89.50). But regardless of the percentage change, this model would be very bad since we would never guess the rarer class. Accuracy is sometimes deceiving.

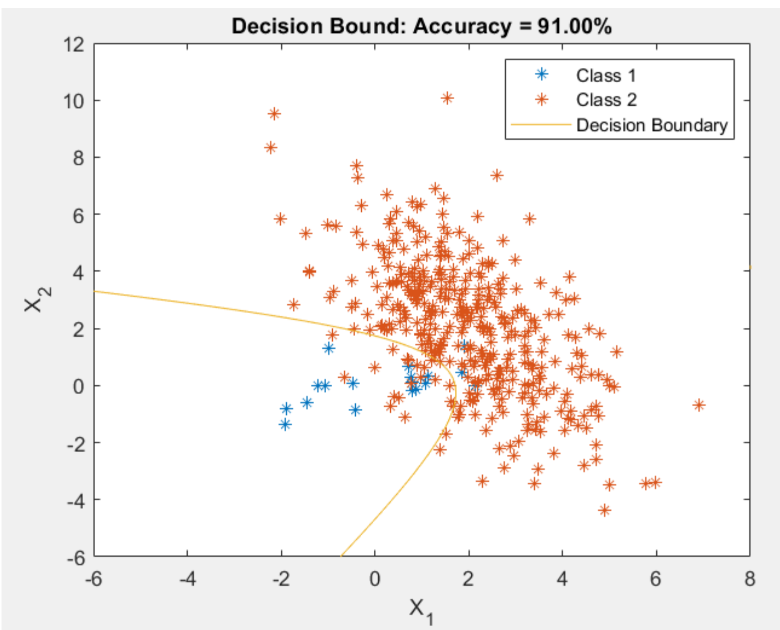
f1) Case 1



f2) Case 2



f3) Case 3



f) There is a significant difference in the plots for part c. The first is linear, while the second and third are nonlinear. The first is also significantly less accurate than Case 3, which is significantly less accurate than Case 2. This is surprising, intuitively one would think that case 3 would have a better prediction as it considers more info than Case 2 (the differing values of sigma). More experimentation & larger data sets would be interesting to see. If we just used the higher prior to make our decision the model would be 95% accurate. This is more accurate than all of the other decision boundaries. Regardless, this model would be very bad since we would never guess the rarer class. Accuracy is sometimes deceiving, especially when one class is much less common.