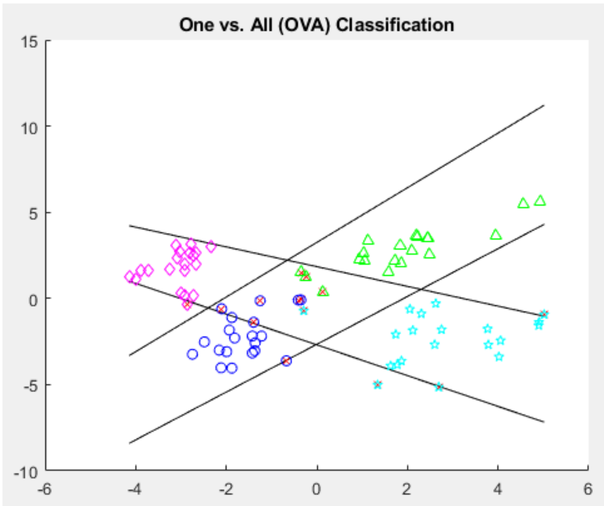
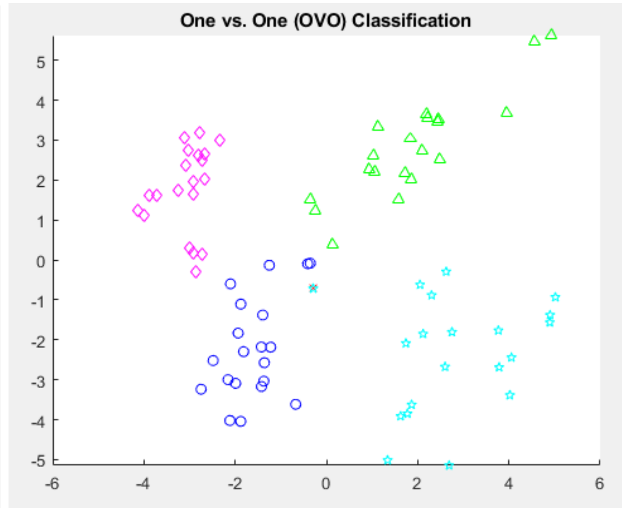
Introduction:
I think this assignment is about comparing the One vs. All (OVA) and One vs. One (OVO) classifications to see which technique has a lower error rate and at the same time is more efficient. By comparing these two, it is important to plot the data with the decision boundaries to not only identify the errors, but also see how much out of bounds room there is where no data point is able to go. In this case, the less out of bounds room, the better, so the areas for each class is larger and more noticeable. For OVO classification, there needs to be a tournament champion associated with each data point where the algorithm will see which class the data point associates with the most based on which side of the boundary line it lies on.

Method:
The method used in this assignment was logistic regression, though OVA used it differently than OVO. This method took in weights starting at 0. For OVA, those weights would be taken in a loop that will go through all of the data points and will choose which class will go against all of the others. If a data point of that class is found, then Y_n is set to 1 and any other class is set to -1. It will then carry out the function sig(Y_n * w^T * x_n) where sig(x) = $e^x/(e^x + 1)$. This is done for each of the data points and they are all added to E_w, which is going to be the final set of weights to show the decision boundary. Being that there are four different classes, there will be four boundary lines to separate the data. Now for OVO, the same algorithm is used, but instead of one class going against all of the other classes, it goes against one other class. So, there are six possible pairs of classes that will be using this algorithm meaning there will be six different boundary lines. After plotting the data and the decision boundaries, the data points that are considered as “errors” are marked with a red ‘x’ to show that the data point is not on the correct side of the decision boundary. Now for finding the OVO tournament champion, every point had to have its own tournament champion. So, for every point, there were six boundary lines causing there to be six votes divided amongst four classes. For every decision boundary, the point could be associated with one of two classes by seeing which side of the boundary line it was on. A class gained a vote if it was on that side and this was repeated for all six boundary lines. Once the six votes were “casted”, the tournament champion would be crowned if it had the most votes. If there was a tie, then the champion would simply be 0. The results are located in the experiments section along with the error rates and images of the graphs.

Experiments:
Error Rates: (ε = 0.0001,   = 0.5)

- OVA: 0.04375
- OVO: 0.0125





Tournament Champions:

4	4	2	2	2	2	2	2	4	2	2	2	2	2	2	2	2	4	2
2	4	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	0
1	1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3															

Class	0 (Tie)	1	2	3	4
Times Won	1	17	17	19	26

Discussion:
As shown in the first graph (OVA), the middle out of bounds section has a lot of data points with the red x's, whereas in the second graph (OVO), there was only one data point that was considered an error (being on the wrong side of the decision boundary). OVO classification appeared to be more accurate, but took longer to compute whereas OVA had a higher error rate, but was very quick to compute. In terms of finding the tournament champions, it seemed like it was divided pretty equally except for class 4 with a lot more tournament wins. There was one tie since a 0 was included in the output meaning that two points had the same number of votes. Overall, it seemed that with a slightly higher runtime, OVO classification was more efficient since it had a much lower error rate than OVA classification.