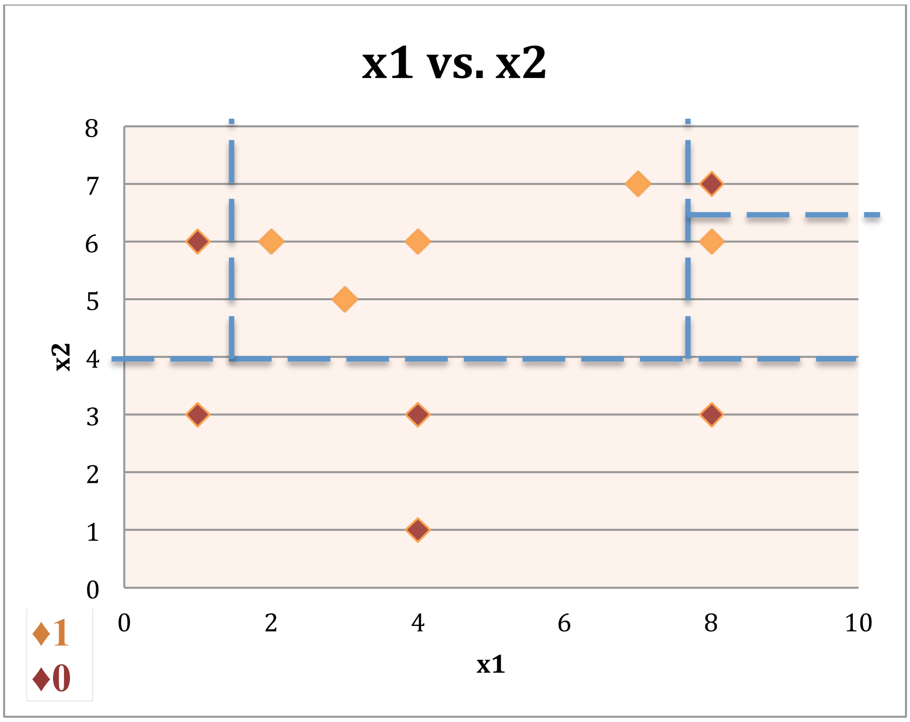
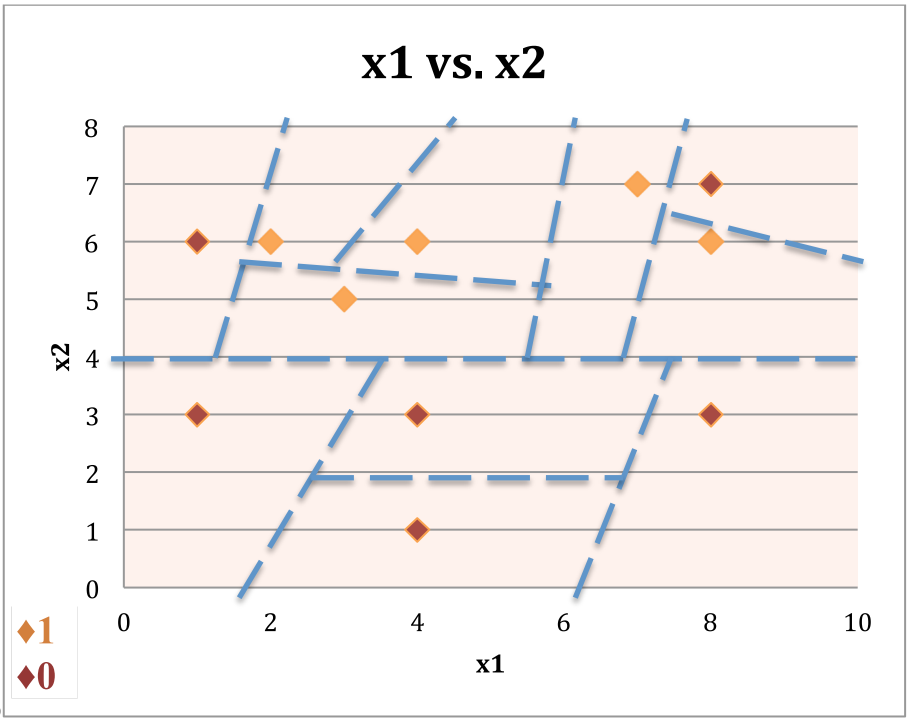
1)

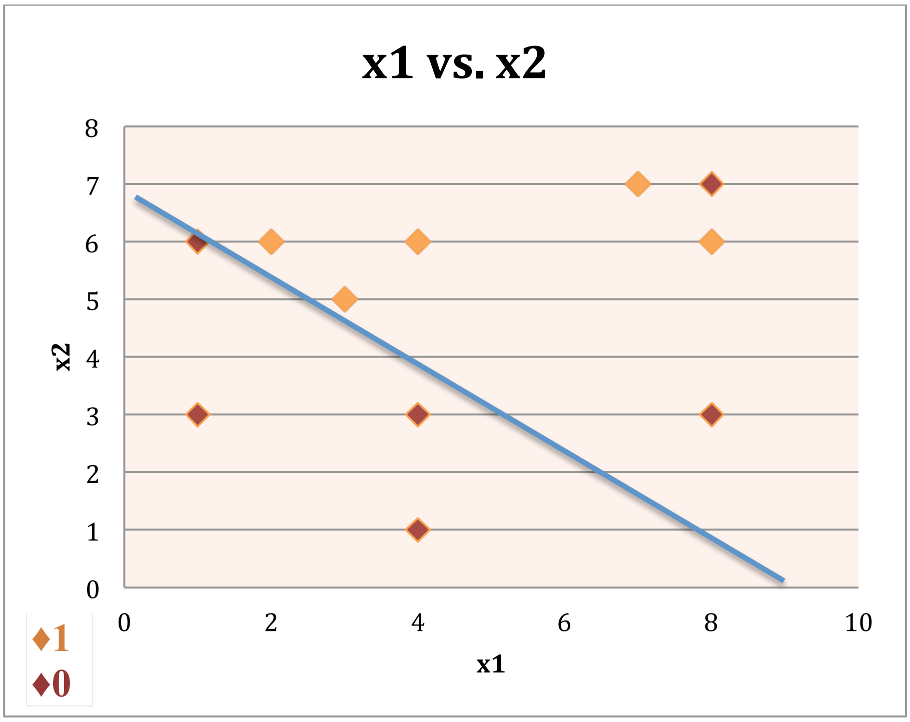
a)

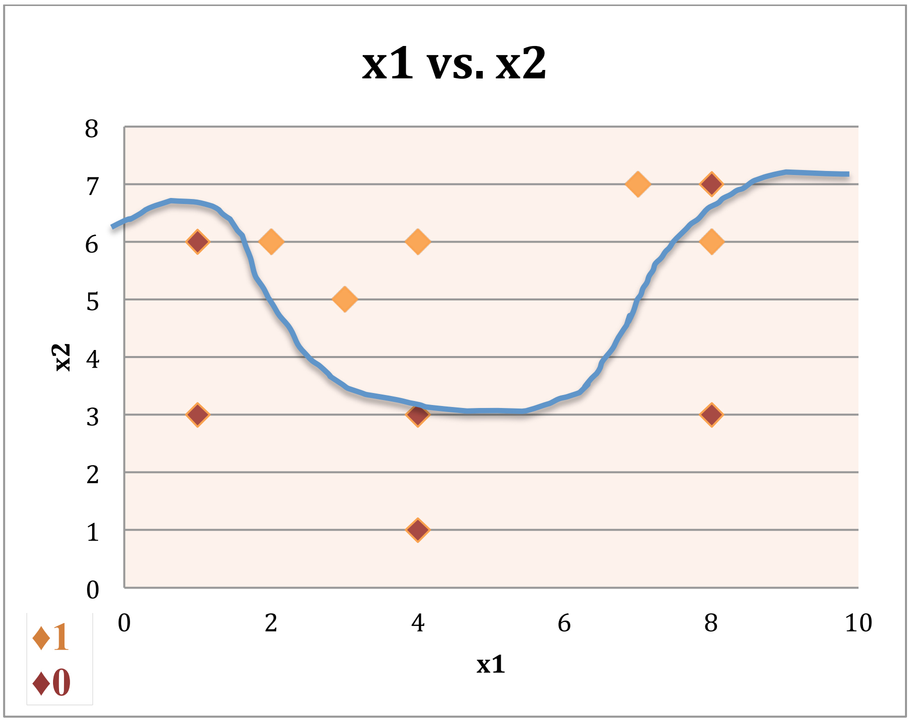

x1 vs. x2
8
7
6
5
4
3
2
1
0
x2
0
2
4
6
8
10
x1
1
0

1)


x1 vs. x2
8
7
6
5
4
3
2
1
0
x2
0
2
4
6
8
10
x1
1
0

2)

3) 

4) 

b) I think the boundary in logistic regression with quadratic terms seems to be the best one fitting. It is because in in decision tree boundaries it there is overfitting, which will result in misclassifications in a possible test set. Also, because there is not a clear clustering the nearest neighbor boundaries will not be helpful in an actual classification. The plain linear regression boundary seems to be nicely fitting as well however, compared to the quadratic one there are more false negatives present.

2)

1) Data: 1, 2, 3, 3, 4, 5, 5, 7, 10, 11, 13, 14, 15, 17, 20, 21

Step 1(random mean choice): $\mu_1 = 1 \,\&\, \mu_2 = 3 \,\&\, \mu_3 = 8$

Step 2(* means it is assigned to c_1, - means it is assigned to c_2):

*1, -2, 3, *3, -4, 5, *5, -7, 10, *11, -13, 14, *15, -17, 20, *21

Step 3 (recalculation):

$$\mu_1 = \frac{1+3+5+11+15+21}{6} = 9.3$$

$$\mu_2 = \frac{2+4+7+13+17}{5} = 8.6$$

$$\mu_3 = \frac{3+5+10+14+20}{5} = 10.4$$

Step 4(test):

$$J(c, \mu) = \sum_{i=1}^{m} \left\| x^{(i)} - \mu_{c^i} \right\|^2$$

$$= \|9.3 - 1\|^2 + \|8.6 - 3\|^2 + \|10.4 - 8\|^2 = 68.89 + 31.36 + 5.76 = 106.01$$