Homework IV

1. Using the following 2-D dataset (x1 and x2 are the attributes and y is the class variable), find the linear SVM classifier. Do your optimization using the dual problem. Namely, provide an explicit expression for the dual optimization problem, solve it (compute the values of the various α_i's) and use the solution to compute the weights attached to the two attributes as well as the bias term.

   [10 Points]

Dataset:

X_1	X_2	Y
-1	1	+
1	-1	-

2. Using the following 2-D dataset (x1 and x2 are the attributes and y is the class variable), find the linear SVM classifier. Do your optimization using the dual problem. Namely, provide an explicit expression for the dual optimization problem, solve it (compute the values of the various α_i's) and use the solution to compute the weights attached to the two attributes as well as the bias term.

   [10 Points]

Dataset:

X_1	X_2	Y
1	0	+
-1	2	-
0	-1	+

3. A SVM is trained with the following data: **[10 Points]**

x_1	x_2	class
-1	-1	-1
1	1	1
0	2	1

Let α_1 , α_2 and α_3 be the lagrangian multipliers for the three data points.

a. Using polynomial kernel of degree 2 what (dual) optimization problem needs to be solved in terms of the lagrangian multipliers in order to determine their values? The polynomial kernel of degree d is given by the equation

$$K(x_i, x_j) = (1 + x_i^T x_j)^d$$

$$where\ x_i\ and\ x_j\ are\ input\ vectors$$

b. Let us say that we have solved the optimization problem and found that $\alpha_1 = \alpha_2 = \frac{1}{8} and\ \alpha_3 = 0$. Moreover b = 0. Can you tell me which of the data points are support vectors. Explain your answer.

c. Assuming $\alpha_1 = \alpha_2 = \frac{1}{8}$, $\alpha_3 = 0$ and b = 0, how will the SVM classify the point (x_1 = -1, x_2 = 0)? Explain your answer?

d. Assuming $\alpha_1 = \alpha_2 = \frac{1}{8}$, $\alpha_3 = 0$ and b = 0, how will the SVM classify the point (x_1 = 1, x_2 = 0)? Explain your answer?

4. Consider the training data given below (Y is the class variable) **[10 Points]**

X	Y
-2	1
-1	-1
1	-1
2	1

a. Assume that you are using linear SVM. Let $\alpha_1, \alpha_2, \alpha_3, and\ \alpha_4$ be the lagrangian multipliers for the four data points. Write the precise expression for the lagrangian dual optimization problem that needs to be solved in order to compute the values of $\alpha_1, \ldots \ldots \alpha_4$ for the data set given above.

b. Do you think, you will get zero training error on this dataset if you use linear SVM? Explain your answer.

c. Now assume that you are using a quadratic kernel $(1 + x_i^T x_j)^2$. Again, let $\alpha_1, \alpha_2, \alpha_3, and\ \alpha_4$ be the lagrangian multipliers for the four data points. Write the precise expression for the lagrangian dual optimization problem that needs to be solved

in order to compute the values of $\alpha_1, \ldots \ldots \alpha_4$ for the data set and the quadratic kernel given above.

d. Do you think you will get zero training error on this data set with quadratic kernel, explain your answer?

5. Support vector machines, like logistic regression models, give a probability distribution over the possible labels given an input example. Explain your answer.

[5 Points]

6. You are trying to use SVMs to build a classifier for a dataset. In the dataset, there are only a few positive training examples and a large number of negative examples. You have to modify the basic SVM dual problem such that none of the positive examples is misclassified but it is ok to misclassify few negative points. Introduce additional parameters and / or constraints in order to achieve this.

[5 Points]