CMPS 142: Homework Assignment 3

Jeffrey Petersen - 1329242
Ben LASTNAME- ID
Raymond Colebaugh - 1377877

May 10, 2015

1.

2. In executing grid search with the specified bounds on parameters, we determined the following costs and gammas for SMO, where X is the cost and y is γ:

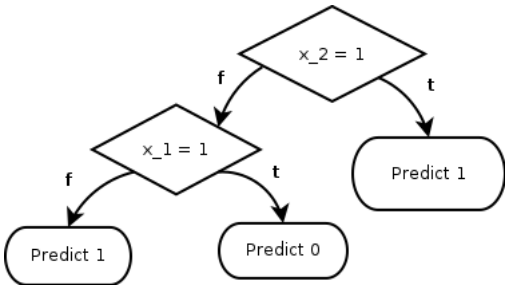
c_{min}	c_{max}	c_{step}	γ_{min}	γ_{max}	γ_{step}	Accuracy
1	16	1	-5	2	1	

   Where the expression for c was i and the expression for γ was 10^i.

3. We can demonstrate the construction of a decision tree on a simple binary dataset such as:

x_1	x_2	y
0	1	1
1	1	1
1	0	0
0	0	1
1	0	0

   In calculating the decision tree, we first decide the attribute to split on for the first node by finding the sum the incorrect predictions. Should we split on x_1, when $x_1 = 0$ there are no errors, and when $x_1 = 1$ there are two errors. Should we split on x_2, if $x_2 = 0$ we have one error, and if $x_2 = 1$ we have no errors. This makes a split on x_2 the better choice. Then for the second node, should $x_2 = 0$, we split on the remaining attribute, x_1. By following decision tree algorithm, we find a representation in the following tree:



   However, there exists a simpler tree to classify this miniscule dataset, should we allow compound conditionals on the nodes: