Kyle Benson
CS 273A - Machine Learning: Fall 2013
Homework 2

Problem 0: Team names

TBD

Problem 1: VC Dimension

(a) We would have $2^d + 2^{d+1} - 2$ parameters, unless you consider that each split node only has 1 parameter due to us only having one feature to choose from, in which case it would only have $2^{d+1} - 1$ parameters.

So the VC dimension would be ??????

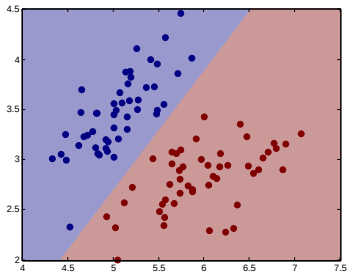
(b) This learner has 4 parameters: the feature index to split on (domain: 1,2), the threshold value (domain: real numbers), and the two choices of which classes to predict (domain: -1, +1).
Example showing that we can shatter 3 points:

blah

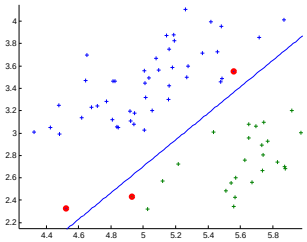
(c) This learner still has 4 parameters, as we've only changed the domain of the feature index to split on.
Example showing how we can now shatter 4 points:

Problem 2: Support Vector Machines

(a) Output thetas (starting with b) are: **-17.2697, 6.3572, -5.3693**

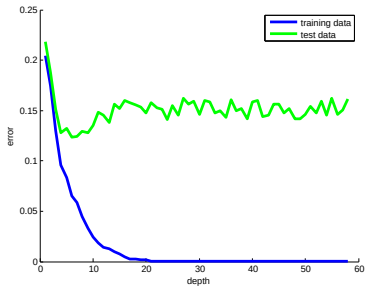


(b) I found the support vectors by extracting the values from the data at the same point the $\alpha_i = 1$. They are circled in red in the plot below.



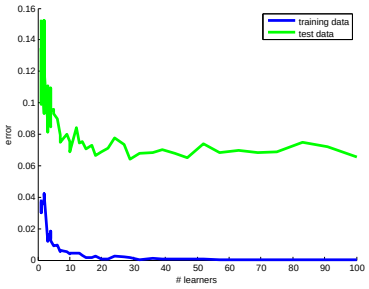
Problem 3: Bagging with Decision Trees

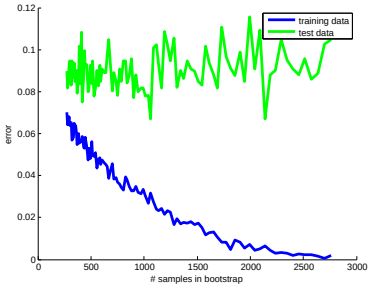
(a) Training error: 3.6219e-04
Test error: 0.1565
Settings: I uses nMin(0), infinite depth, minScore(.001), and all available features for splitting.



(b)

(c) We should believe that we are overfitting with these parameters because they introduce more complexity and, for example, learn training data completely with deep enough trees.





(d) In this figure we see that

Problem 4: Boosting

(a) TBD