Daniel Campos October 17st, 2013

RCS: Campod2 RIN:660996361

Machine Learning CS 4100

I worked on these problems with Zoe Konrad☺

Problem Set 8

1. Exercise 4.3 (200)
   1. When complexity of F increases the deterministic noise will go up and there will be a higher tendency to over fit.
   2. When complexity of H decreases the deterministic noise will go up and there will be a lower tendency to over fit since the hypothesis set is less able to
2. Exercise 4.5 (200)
   1. For this one we want  is an identity matrix because that will cause the result to be the sum squared



* 1. For this one we want  to be equal to a matrix of all 1 s multiplied by a constant (see attached paper for the proof)

1. Exercise 4.6 (100)

Hard order constraints are better than a soft order constraint because in the perceptron of binary classification a soft order constraint will do nothing. Hard order constraints the weights can be scaled down by  and still produce a accurate classification for the model. The model may not represent all the points of the data set but it will work for most and be useful for regularization

1. Exercise 4.7 (200)
   1. Since  and  and given 

🡪 🡪

🡪



* 1.  Based on part a

 since only when the g(x) is not equal to the y contributes to the variance

thus



* 1. Assuming I was correct on previous part I use the notion that error will never be greater than .5 because if it is then we just flip the classification for the learned g. Based on this, any g-

🡪

🡪

🡪



* 1. There will be no upper bound for error of data because we just flip it if we have to.
  2. When we train with fewer points I expect  to be higher since there are less points to make the model work and we are more likely to over fit to the points we have.
  3. By increasing our validation set we can have a lower Eout if we have a drop in  however if we increase the validation set too much we will not have enough to train on and thus the  will not be at all useful and thus our Eout will suffer.

1. Exercise 4.8 (100)

Yes it is an unbiased since the choice of any  was not influenced at all by the data. It has been separate and thus is unbiased o the data. The data is not changing for every hypothesis set we run and thus is unbiased.