Daniel Campos September 18, 13

RCS:Campod2 RIN:660996361

CSCI 4100 Machine Learning

I worked on the HW with Zoe Konrad

Problem Set 3

1. Exercise 1.13 (100)
   1. The probability of error is equal to 

Which is 

* 1. The performance of h is independent of  when  because at that value half or less the Y values are related to f(x) and therefore the hypothesis performance will be independent of 

1. Exercise 2.1 (100)
   1. 







Breakpoint at 2

* 1. 









breakpoint at 3

* 1. 



No breakpoint since  is impossible

1. Exercise 2.2 (100)
   1. Verify the previous
      1. K=2



 So it is correct

* + 1. K=3







So it is correct

* + 1. Does not apply since that equation had no breakpoints
  1. There is not a hypothesis set since the case is faster than a polynomial though is short of 2^n and by the definition is must not be faster than a polynomial

1. Exercise 2.3 (100)
   1. 1. 
      2. 
      3. 
2. Exercise 2.6 (100)
   1. Given 🡪 

🡪 





The error for Ein is 0.115090370633255

The error for Etest is 

 

the error for e test is 0.096032279131967

Where 1 replaces  because we only have 1 hypothesis, G in this case the error is 0.0588875

* 1. Yes there are reasons you should reserve less examples for testing. This would occurs when you do not have enough example to learn on and therefore by having a low number of learning data, though your test set would approximate your normal function, you hypothesis would likely be wrong. It is a careful game of keeping the Etest close to the Eout and making sure that the G from the Ein is actually relevant and probably right.

1. Problem 1.11 (200)

Assume based on y=+1 and y=-1

For the CIA Fingerprint recognition

False Reject (y=+1)



False Accept(y=-1)



For the Supermarket

False reject(y=+1)



False Accept(y=-1)



1. Problem 1.12 (300)
   1. 
      1. First we rewrite  as 
      2. 
      3. 

where the last term is always positive and minimized at h is hmean. The first term is the sum of squared derivation about sample mean

* + 1.  which by definition of hmin =0
    2. so Ein(h) is minimized when h=h mean
  1. Proof by induction

 for all point n>2

Now assuming the minimum is achieved when all possible sets of h where n>2. For points n=1 the median will be the only value and for two there is no median

Consider case n+2 points. is the minimum and  is the maximum

 is minimized for any value when

 Minimize the sum of the remaining point for n point and by our inductive hypothesis this will occur at their median, which will be the same median as the n+2 points since we have removed two points.

Example

Consider data with values {1,1,3,3,4,6,9} our median value is 3. Our absolute derivation is {2,2,0,0,1,3,6} which in turn has a median value of 1 so the absolute derivation is 1.

* 1. Since only a single point is perturbed, the H Median will not change(unless the perturbed point was the median or smaller). This will not change because regardless of how big one of the points the values change 50% of the remaining values are large and 50% are smaller. If the median were this perturbed value or smaller our median would shift 1 down in order to make the remaining values be 50% bigger and 50% smaller. For the Mean, this high value will cause our mean to be high and be unreliable and unusable. Medians are best used to deal with outliers since they are mostly unaffected by outlier or affected less than other statistical types.