Lecture - 7 - Mx869t

Taken :- Mx869t

Summary :- We started with VC analysis where we got the relation of generalization and approximation with respect to the d_{VC}.

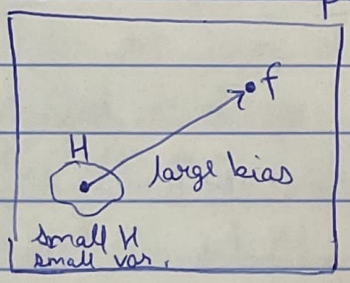
$d_{VC} \uparrow \Rightarrow E_{in} \approx 0$ (Higher chance of approximating)

$d_{VC} \downarrow \Rightarrow E_{in} \approx E_{out}$ (Higher chance of generalising)

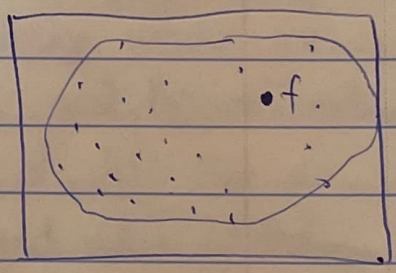
Then we studied about Bias-Variance. This is an alternate view of the approximation-generalization tradeoff. Where bias means how well can hypothesis approximate to our target function f?

& Variance means How close can we get to this using a finite dataset?

Consider Example.



Here only 1 hypothesis is present. So small var, but the hypothesis is far from the target function that is why we have a large bias.



Here we have Large Hypothesis set but close to the target function. That is why small bias & large Var.

then we studied Least squares Linear regression where

$$\text{error}\left(h(x), f(x)\right) = \left(h(x) - f(x)\right)^2.$$

($h(x)$: prediction; $f(x)$: actual.)

Later we looked upon a simple learning problem for function $f(x) = \sin \pi x$. Where we studied the difference between the flat lines and All lines. Where we saw the results that for higher number of data points All lines is better and for small number of data points flat lines tend to give better performance.

We speculated one Error formula

$$E_{out} = \text{bias} + \text{var}$$

Then we studied the graph of the learning curve for simple model and complex model, where we analyzed the trends of ~~Error~~ Eout and Ein.

And at the end we took a ~~glance~~ look at the pocket algorithm and the digits data.