Northeastern University
College of Engineering
Group member name(s): Yanming Liu
Group member UID(s): 002199402

May 27, 2022
Assignment 5

IE 7374: Machine Learning

Prove the bias-variance decomposition below:

$$\iiint_{Dxy} ((h_D(x)-y)^2 p(x,y)p(D))\,dx\,dy\,dz = E_{x,D}[(\overline{h}(x)-h_D(x))^2]+E_x[(\overline{h}(x)-\overline{y}(x))^2]+E_{x,y}[(\overline{y}(x)-y(x))^2] \tag{1}$$

Answer to Question

The left part in formula (1) can be represented as follows,

$$\iiint_{Dxy} ((h_D(x)-y)^2 p(x,y)p(D))\,dx\,dy\,dz = E_{D,x,y}[(h_D(x)-y)^2]$$

Let $y(x) = \overline{y}(x) + \varepsilon$, ε is noise with a mean of 0, i.e.,$E(\varepsilon) = 0$

$$\begin{aligned} E_{D,x,y}[(h_D(x)-y)^2] &= E_{D,x,y}[(h_D(x)-\overline{y}(x)-\varepsilon)^2] \\ &= E_{D,x,y}[(h_D(x)-\overline{y}(x))^2 + 2\varepsilon(h_D(x)-\overline{y}(x)) + \varepsilon^2)] \\ &= E_{D,x,y}[(h_D(x)-\overline{y}(x))^2] + E_{x,y}[\varepsilon^2] \\ &= E_{D,x}[(h_D(x)-\overline{y}(x))^2] + E_{x,y}[(y(x)-\overline{y}(x))^2] \end{aligned} \tag{2}$$

Then, we modify the first item in the right part in formula (2),

$$\begin{aligned} E_{D,x}[(h_D(x)-\overline{y}(x))^2] &= E_{D,x}[(h_D(x)-\overline{h}(x)+\overline{h}(x)-\overline{y}(x))^2] \\ &= E_{D,x}[(h_D(x)-\overline{h}(x))^2] + E_x[(\overline{h}(x)-\overline{y}(x))^2] \\ &\quad + 2E_{D,x}[(h_D(x)-\overline{h}(x))(\overline{h}(x)-\overline{y}(x))] \\ &= E_{D,x}[(h_D(x)-\overline{h}(x))^2] + E_x[(\overline{h}(x)-\overline{y}(x))^2] \end{aligned} \tag{3}$$

The reason why we get formula (3) is $E[h_D(x)] = E[\overline{h}(x)]$, using (3) to substitute (2),

$$E_{D,x,y}[(h_D(x)-y)^2] = E_{x,D}[(\overline{h}(x)-h_D(x))^2] + E_x[(\overline{h}(x)-\overline{y}(x))^2] + E_{x,y}[(\overline{y}(x)-y(x))^2]$$

Prove.